MAKING SENSE OF GENOTYPE x ENVIRONMENT INTERACTION OF *Pinus radiata* IN NEW ZEALAND

A thesis submitted in partial fulfilment of the requirements for the

Degree of Master of Forestry Science

at the

University of Canterbury

by

Timothy Myles McDonald

B.For.Sc. (Hons), B.Sc., M.N.Z.I.F.

School of Forestry

University of Canterbury

2009

ABSTRACT

In New Zealand, a formal tree improvement and breeding programme for *Pinus radiata* (D.Don) commenced in 1952. A countrywide series of progeny trials was progressively established on over seventy sites, and is managed by the Radiata Pine Breeding Company (RPBC). Diameter at breast height data from the series were used to investigate genotype x environment interaction with a view to establishing the need for partitioning breeding and deployment efforts for *P. radiata*. Nearly 300,000 measurements made this study one of the largest for genotype x environment interaction ever done.

Bivariate analyses were conducted between all pairs of sites to determine genetic correlations between sites. Genetic correlations were used to construct a proximity matrix by subtracting each correlation from unity. The process of constructing the matrix highlighted issues of low connectivity between sites; whereby meaningful correlations between sites were established with just 5 % of the pairs. However, nearly two-thirds of these genetic correlations were between -1.0 and 0.6, indicating the presence of strong genotype x environment interactions.

A technique known as multiple regression on resemblance matrices was carried out by regressing a number of environmental correlation matrices on the diameter at breast height correlation matrix. Genotype x environment interactions were found to be driven by extreme maximum temperatures (*t*-statistic of 2.03 against critical *t*-value of 1.96 at 95 % confidence level). When tested on its own, altitude was significant with genetic correlations between sites at the 90 % confidence level (*t*-statistic of 1.92 against critical *t*-value of 1.645).

In addition, a method from Graph Theory using proximity thresholds was utilised as a form of clustering. However, this study highlighted the existence of high internal cohesion within trial series, and high external isolation between trial series. That is, grouping of sites (in terms of diameter) was observed to be a reflection of the series of trials for which each site was established. This characteristic is particularly unhelpful for partitioning sites into regions of similar propensity to genotype x environment

interaction, as the genotype x environment effect is effectively over-ridden by the genotype effect.

Better cohesion between past, present and future trial series, and more accurate bioclimatic data should allow more useful groupings of sites to be extracted from the data. Given this, however, it is clear that there are a large number of interactive families contained in the RPBC dataset. It is concluded that partitioning of New Zealand's *P. radiata* breeding programme cannot be ruled out as an advantageous option.

Key words: *genotype x environment interaction, regionalisation, proximity matrix, multiple regression on matrices, Graph Theory, proximity thresholds*

ACKNOWLEDGEMENTS

I would like to thank the Radiata Pine Breeding Company for the opportunity to conduct this research and for the support that has been given to the project. In particular, thanks for Paul Jefferson and Toby Stovold for assistance with obtaining, collating, and sorting the data.

Dr. Luis Apiolaza has been a very proactive supervisor, particularly in the early stages of my thesis, enabling a productive start to the project. He was then able to bail me out when I encountered some unexplained programming issues. Without this, the project would not have progressed – muchas gracias Luis.

Special thanks also to Dr. Marco Reale, Dr, Carl Scarrott, and especially Professor Mike Steel, from the University of Canterbury's Department of Mathematics and Statistics, for their time and suggestions. Also, thanks to Associate Professor Richard Woollons for his analytical guidance and discussions about forestry in general.

In addition, I am very appreciative of the generous assistance given by Dr. Hamish Cochrane with GIS issues, the ever-vigilant manuscript-reviews performed by Associate Professor Euan Mason, and the tireless administrative efforts of Jeanette Allen and the School of Forestry technicians: Dave Clark, Lachlan Kirk, Nigel Pink, and Vicky Wilton.

A big thanks to Mum for getting me to where I am and supporting my various endeavours. Thanks to Justin, Alwyn, Yvette, Norbert, and Timo for entertainment in-between the moments of analysis and writing, and for tips on all things post-grad. Canterbury Rowing Club & University of Canterbury Rowing Club for providing me with a well-needed distraction from academia. In particular to "Doc", Chipper, and Raf for all the miles we've done together on Kerrs Reach – Seniores Priores.

Finally, vielen dank Kerstin. Being in a relationship with a post-graduate student is not much of a life at times. Thanks for being patient. Thanks for looking after me. Our roles reverse next year, and I'll do my best to do as good a job as you have.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ASReml	Statistical software for fitting linear mixed models using residual maximum likelihood.
ASReml-R	Statistical software for fitting linear mixed models using residual maximum likelihood in the R-environment.
CliDB	National Climate Database
CliFlo	A web-based service for provided by NIWA for accessing CliDB
GPS	Geographic Positioning System
GxE	Genotype x Environment Interaction
MDS	Multi-Dimensional Scaling
MRM	Multiple Regression on Resemblance Matrices
NIWA	National Institute of Water and Atmospheric Research
R	A language and environment for statistical computing and graphics
RPBC	Radiata Pine Breeding Company Ltd.
SAS	Statistical Analysis Software

1. INTRODUCTION

1.1 General introduction

Forest tree improvement is the practice of understanding and exploiting variation, using principles from genetics, statistics, economics and reproductive biology, to produce more profitable or desirable trees. By controlling breeding and prudently selecting parents, tree growth and quality can be changed and often improved (Zobel and Talbert, 1984). These tree-breeding skills are most effective when combined with silvicultural skills that help manipulate the environment in which a tree is growing, resulting in production of superior phenotypes.

In general, a genotype is considered to be superior in commercial forestry when it outperforms other genotypes with respect to a trait that is of economic interest, or in some cases aesthetic appeal. Traits of forest tree species are known to have greater variability than species of many other organisms (Zobel and Talbert, 1984). Without this variation within species, attempts to improve forest trees through genetics would most likely prove unsuccessful. In fact, it is this variation from one phenotype to the next which enables superior trees to be produced.

The variation described above is, in part, due to what is known as genotype x environment interaction. Genotype x environment interactions bring about changes in the relative performances of genotypes when grown in two or more different environments, such that a genotype that is considered outstanding in one environment may be considered ordinary (or less than ordinary) in a second environment.

By gaining greater insight into genotype x environment interactions of *Pinus radiata*, it is hoped that breeders will be able to more accurately predict the performance of genotypes across the range of New Zealand environments. Consequently, the probability of selecting the most suitable parent, or group of parents, for any site will be improved. Furthermore, an understanding of the likely performance of genotypes on a range of sites will enable a decision to be made regarding the breeding strategy for New Zealand. Is it to the forestry industry's advantage to create regional breeds or

is it more efficient to produce a national breed from a central location and deploy it across the country?

In addition, a more comprehensive knowledge of genotype x environment interactions allows forestry companies to be more flexible. Forest managers can respond quickly to changes in breeding objectives by allocating to sites those genotypes that are considered more appropriate to meet the new objectives. Should genotype x environment interactions be shown to be significant, forest managers will also be in a better position to respond to changes in rates of reafforestation, through the maintenance of larger seed orchards (Shelbourne et al., 1986). As local demand for seed increases, it can be sourced internally. Conversely, should local demand decrease, forest managers can export surplus seed to areas with similar climatic attributes confident that seed will respond in a manner comparable to its performance on local sites.

From a more global perspective, an increased population is expected to bring about a shortage of prime land and an increase in demand for forest products (Zobel and Talbert, 1984). This will require both the development of breeds that are successful on marginal and sub-marginal land, and an increased productivity of breeds that are grown on fertile land.

There are many end-uses for which *P. radiata* is grown, and the species has been noted as having large phenotypic plasticity compared to many other tree species (Jayawickrama et al., 1997, Thulin, 1957). Early anticipation from Thulin (1957) was that, despite this variation, specific geographical strains of *P. radiata* could not be expected due to its confined and uniform natural habitat in California. However, native provenances have since been recognised and it has been suggested that there are distinct advantages to be gained from provenance material, particularly in terms of edaphic tolerances, wood properties, and potential for hybridisation (Burdon et al., 1997, Shelbourne et al., 1986).

It is advantageous to utilise the variation which is inherent in *P. radiata*, and understanding genotype x environment interactions can facilitate this goal. From an economic perspective, forestry companies wish to move to shorter rotations: earlier

harvest ages have the effect of not only bringing forward cash inflows (i.e. sales), but also decreasing the periods that many cash outflows (e.g., land preparation, establishment, and silviculture) are carried for. However, in recent years, this author has observed the larger New Zealand forestry companies extending rotation ages in an effort to improve certain wood quality characteristics (i.e. density and stiffness): this strategy has been noted by others in New Zealand forestry (Walker, 2007, Sutton, 2007). The tactic of increasing rotation ages to alleviate lower than desired wood quality characteristics can act as merely a short-term fix at best. The New Zealand forestry industry must not preclude the development of higher density (or other wood quality characteristic) trees with short rotation lengths. Exploitation of large tree-to-tree variation provides a solution to the decline in quality of those traits, especially in combination with a tendency for wood quality characteristics to show strong heritability and low genotype x environment interactions (Zobel and Talbert, 1984).

1.2 Statement of the problem

Growth and quality of genotypes are affected by environmental variables, such that the performance of a genotype (relative to its peers) in one environment may not reflect the performance of that same genotype in a different environment. Such interactions need to be understood to ensure that breeding programmes are more accurate in their prediction of genetic gain.

In New Zealand, partial attempts have been made to explain genotype x environment interactions for *P. radiata* (Burdon, 1977, Johnson and Burdon, 1990, Carson, 1991). However, these studies have been limited by the scale of the data studied. For this thesis, the intention was to explain the genotype x environment interactions observed from measurements made from a combination of over seventy New Zealand sites and more than 2,500 distinct genotypes.

1.3 Background

1.3.1 Origin of G x E data

Breeding programmes for conifers began around the world after the end of World War II (Shelbourne et al., 1989). In New Zealand, a formal tree improvement and breeding programme commenced in 1952, although it was limited in terms of sites and provenances due to earlier assumptions that land race material would be sufficient as the gene resource population, and that differences between provenances were insignificant (Shelbourne et al., 1986, Burdon et al., 1997). However, it was thought that this programme would eventually be able to provide information on the effect of environment on genotype performance (Thulin, 1957). The programme became progressively more comprehensive through the evolution of tree improvement theory, methodology, and technology, leading to the establishment of a countrywide series of provenance tests during the 1980s on a large number of sites (Burdon et al., 1997).

In the mid 1980s, with the restructuring of the New Zealand Forest Service, the responsibility for tree improvement work shifted to the private sector. In response, the New Zealand Radiata Pine Breeding Co-operative was formed between the Forest Research Institute and nine forestry organisations (Jayawickrama et al., 1997). The organisation has since expanded, with several major New Zealand and Australian companies as shareholders, and is now known as the Radiata Pine Breeding Company (RPBC) Ltd. The RPBC aims to provide superior radiata pine to its shareholders and customers in Australasia[1]. The area owned/managed by the RPBC shareholders encompasses a vast number of heterogeneous environmental zones.

Since the inception of New Zealand's tree improvement programme for *P. radiata*, over 2650 trees have been selected (based on general combining ability), and progeny tested (Jayawickrama et al., 1997). It should be noted that individuals were phenotypically selected, based on external traits such as height, diameter, straightness, lack of malformation, general health, small branch size, and freedom from stem cones. Additional breeds were later developed, as so-called 'elite' populations (with

[1] http://www.rpbc.co.nz/ accessed 25th September 2007 at 10:45am

the aim of improving intrinsic properties like spiral-grain and stiffness), and matched to sites where they were likely to be of most value.

Part of the RPBC's programme involved the establishment and measurement (mostly between the ages of 7-10 years old) of more than seventy progeny trials across the length of New Zealand (Figure 1 and Appendix A). These trials were populated by more than 2,500 genotypes, making available stem growth and form data for approximately 0.3 million trees. In addition, wood quality data has been collected for a subset of those trees.

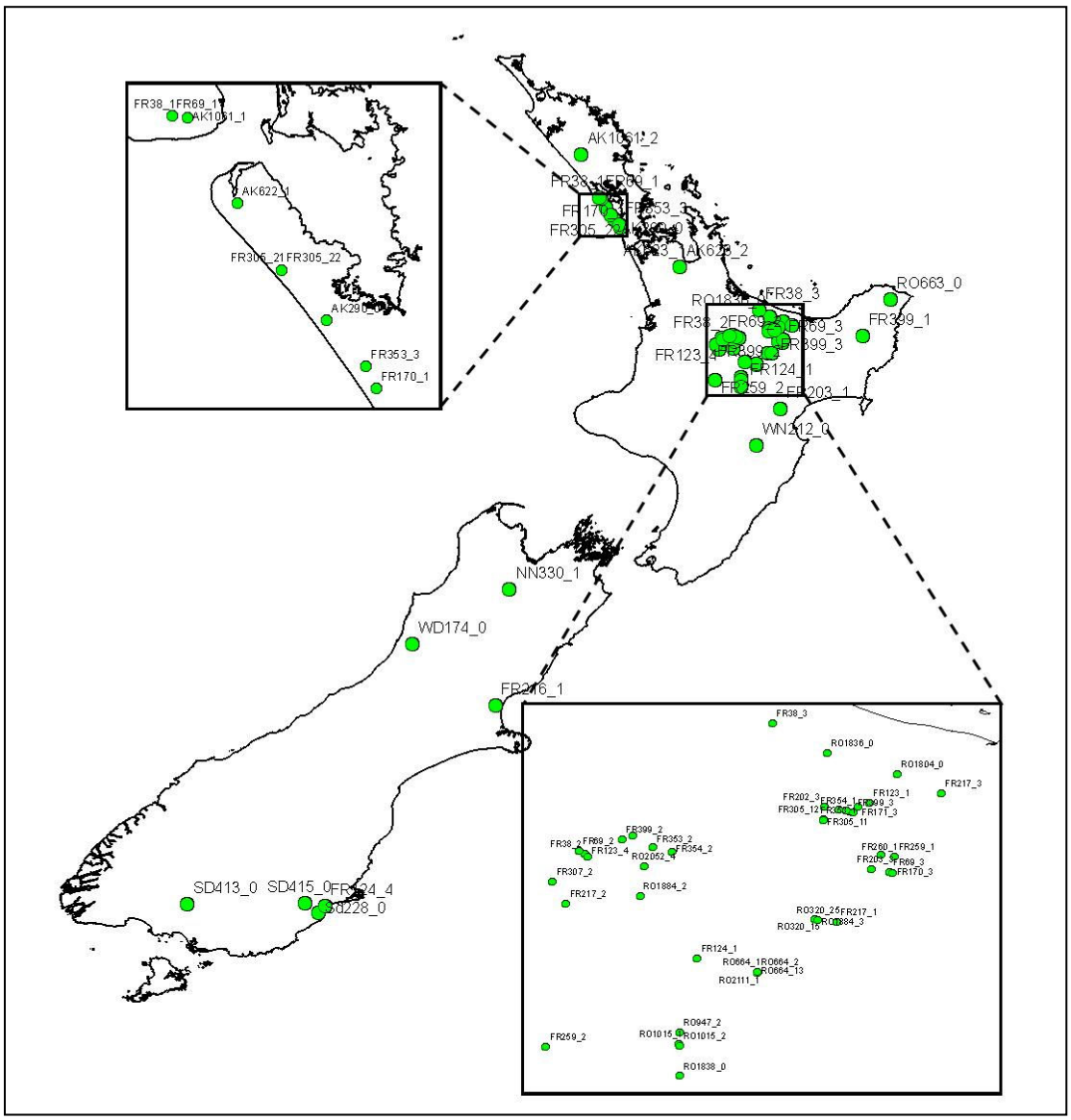


Figure 1: Location of RPBC progeny trials in New Zealand.

1.3.2 G x E interaction

Genotype x environment interaction is a term used to describe the interaction of environmental factors and genes (or particular sets of genes). It is often the case that there is a change in the performance ranking of a given genotype when grown in different environments, although relative performance of genotypes may change with no change in ranking (Zobel and Talbert, 1984). Genotype x environment interaction can be described as a term in a generic additive model for the phenotype of an individual:

$$P = \mu + G + E + GxE + \varepsilon \qquad [1]$$

where P is the phenotypic expression of a trait, μ is the population mean for that trait, G is the effect of the genotype, E is the effect of the environment, GxE is the interaction of the genotype and the environment acting on the expression of the trait, and ε is an error term.

In such an interaction, the expression of a trait (e.g. diameter growth) in one environment can be thought of as a different characteristic to the expression of that same trait in a different environment (Falconer, 1952). The performance of a genotype in the first environment is, then, genetically correlated to the performance of that genotype in the second environment. This allows for the performance of a genotype in two environments to be studied using genetic correlation.

There has been much research into genotype x environment interaction (Amer et al., 1992, Falconer, 1952, Mathur and Horst, 1994, Mulder and Bijma, 2005, Simons and Roff, 1996, Yamada, 1962). However, this research has tended to focus on agricultural crops or animals. In the case of agricultural crops, weather in one year can be vastly different to the next, making the effect of the environment random (Burdon, 1977). In a forest environment, changes in weather over a rotation are averaged-out: effectively, making the environment a fixed effect. In the case of animals, the same individual is able to be observed in two or more environments. This is not so with tree species, where an individual can only be observed on one site (although the situation may change in the near future with a clonal approach), and its family must be used to assess its (likely) individual performance on a range of sites (Shelbourne et al., 1986).

Where trees, or forests, have been studied, much of this research has been outside New Zealand (Ades and Garnier-Gere, 1997, Hodge and White, 1992, McKimmy and Campbell, 1982, Owino, 1977b, Owino, 1977a, Owino et al., 1977, Owino and Zobel, 1977b, Owino and Zobel, 1977a). In an Australian study, Matheson and Raymond (1984) reported that all *P. radiata* families tested changed rank across sites. In addition, ranking reversals occurred from one site to another: indicating considerable genotype x environment interaction had occurred. Matheson and Raymond (1984) did not attempt to explain the observed genotype x environment interaction, only to assess the significance of such interactions to breeding *P. radiata* in Australia.

Where research has been conducted on *P. radiata* in New Zealand, little has been concluded about the effect of genotype x environment interactions on either growth or wood quality (Burdon, 1975, Carson, 1991, Johnson and Burdon, 1990). Where promising results have been found (Kumar, 2004, Kumar et al., 2008), applications of these have been limited by the number of sites and/or genotypes involved in the studies. To date, there is no published evidence showing selected *P. radiata* genotypes with congruent performances in all, or most, environments (Matheson and Raymond, 1984), although genetic rankings for many wood properties appear to be more stable than those for growth (Kumar, 2004).

From an applied perspective, Zobel and Talbert have labelled the necessity of establishing trees in vastly different environments as “the trademark and challenge of forestry” (1984, p57). This is increasingly so with a declining productive forest-land base requiring establishment of forest stands on sites that are considered sub-optimal for tree growth. Intensive land preparation, a matter-of-course in New Zealand plantations, and fertilisation may also have profound effects on the environment in which trees are to be established; especially if there is disparity between the improved site and the site-type that improved stock were developed to grow on.

It has been recognised that genotype x environment interactions are the likely reason that superior genotypes in a particular environment are not necessarily superior in all other environments (Burdon, 1977). Jinks and Mather (1955) showed the interaction of a genotype with its environment to be a highly heritable character. Therefore, if the performance of a genotype is unpredictable from one environment to the next, its

inclusion in a breeding programme is undesirable, as expected gains are difficult to predict (Owino, 1977b). However, genotypes that show stable performance across environments are not necessarily desirable either as they are unable to take full advantage of improved environmental conditions.

Owino (1977b) suggests that it is imperative to have genotypes with superior volume production and adaptability to a wide range of sites. This approach, however, appears to neglect the importance of wood quality traits. Burdon (1977) posed the question of whether tree breeding should focus on producing genotypes suitable for specific environments or genotypes suited to a wide range of environments. Certainly, if one genotype, or a group of genotypes, perform well over a variety of environments, it would be feasible to select only for those genotypes (Zobel and Talbert, 1984). A similar idea has been mooted in animal breeding (Mulder and Bijma, 2005). The solution is dependent on the practical importance of genotype x environment interactions (Matheson and Raymond, 1984). Where interactions are of little importance, breeding programmes can focus on a few genotypes suited to a wide range of sites. In contrast, if interactions heavily affect phenotypic responses, environmental preferences must be taken into consideration.

A common method of assessing the practical importance of genotype x environment interactions is to use a measure of the loss of potential genetic gain (Matheson and Raymond, 1984). However, genetic gain prediction is based on selection theory, which assumes there is no genotype x environment interaction (Owino, 1977a). Owino *et al*. (1977) found that bias caused by genotype x environment interaction resulted in predicted genetic gains being twice as much as they should be. Genotype x environment interaction has also been found to affect accuracy of selection, selection intensity, genetic variance of the breeding goal, and heritability (Mulder and Bijma, 2005, Zobel and Talbert, 1984).

1.3.3 Regionalisation[2]

Breeders facing genotype x environment interactions have several options from which they must choose their breeding strategy:

- breed for a single site;
- breed for an average environment;
- exclude highly interacting genotypes; and
- group environments into breeding regions and select groups of genotypes that are best adapted to particular regions.

The first option does not address the variation caused by genotype x environment interactions. Both the second and third options are simple solutions to implement, but for any one environment will theoretically yield modest amounts of genetic gain (Ades and Garnier-Gere, 1997). The fourth option requires a more complex breeding programme, but should produce larger genetic gains. It is, however, more expensive to implement. In such an approach, care must also be taken to ensure regions do not become too small, as this will lead to many inefficiencies (Zobel and Talbert, 1984).

Shelbourne *et al*. (1986) list the issue of whether or not to regionalise as a top-five priority for research into New Zealand's breeding programme. However, a single *P. radiata* breeding programme has been, and continues to be, used for the whole country, owing mostly to doubts concerning benefits to be gained from regionalisation (Carson et al., 1990, Carson, 1991). The general adaptability of various *P. radiata* provenances may also have had some influence on the continuance of this strategy. In particular, the Monterey provenance has been found to be tolerant of many different soil types and most New Zealand conditions except colder southern sites and those with snow, where the Año Nuevo provenance has been found to perform well (Burdon et al., 1997). In addition, genotype x environment interaction is difficult to assess. There may be a significant effect for one trait, but not for another (Zobel and Talbert, 1984). In this case, the relative economic weighting of the affected trait to the breeding objective will be important in determining the importance of the interaction.

[2] The term regionalisation does not imply contiguous geographical areas, but rather a number of sets of areas for which selected genotypes will be the same.

It may make more sense to utilise regional orchards to produce special-purpose breeds than to subdivide the breeding population itself (Shelbourne et al., 1986).

The decision of which breeding strategy to choose will depend on the level of genotype x environment interaction as well as the cost of alternatives. It is important that resources are not allocated to a regional strategy unless there is an identified and significant partitioning of breeding zones (Ades and Garnier-Gere, 1997). Johnson and Burdon (1990), in a study of *P. radiata* in the New Zealand region of Northland, were able to select families for which regionalisation improved genetic gain to just 25 % as compared to 22 % from non-regionalisation. Carson (1991) also found only a slight increase in genetic gain for diameter through regionalisation of seed orchards: 11.2 % for a non-regionalised programme compared to 14.4 % for a programme with 11 regions.

If the level of interaction is not large, for the same expenditure, a breeder may generate greater gains from a non-regionalised programme because selection intensities may be increased with a larger breeding population (Carson, 1991). In contrast, high levels of interaction will necessitate separate breeding populations for different regions. In this case, smaller breeding populations will be required in each region if the total resource allocation for breeding remains the same, due to increased implementation and progeny testing costs.

In New Zealand, a prime candidate for establishing a regional breed is the Northland clays region. This region has poor genetic correlations with the central pumice plateau, where most progeny testing is currently done, due to clay soils being phosphorus deficient (Burdon, 1971). As discussed previously, Johnson and Burdon's (1990) findings do not justify Northland clays having its own breeding population. And, if regionalisation is not worthwhile for the Northland clays, it may not be worthwhile elsewhere in New Zealand where genetic correlations with the central pumice plateau are stronger (Johnson and Burdon, 1990). Several New Zealand-based studies agree with this result. Shelbourne and Low (1980) found genotype x environment interaction not strong enough or sufficiently pronounced with a regional pattern to warrant regional breeding programmes. Carson (1991) was able to achieve 90% of the maximum possible genetic gain when selecting from only the best site.

Where there was more than one site to a region, the overall gain was just 0.6% of breast height diameter at age 9 which is unlikely to cover the increased cost of regionalisation. It should be noted, however, that Johnson and Burdon's (1990) study included only four sites and Shelbourne and Low (1980) analysed just five sites. Carson (1991) reported findings from eleven sites, but these progeny tests utilised just 25 parents. As a result, the conclusions drawn from these studies are probably insufficient to dispel the need for regionalisation.

Achieving a reduction in the ratio of interaction to family variance components for diameter, led Matheson and Raymond (1984) to suggest that instead of regionalising breeding programmes in Australia, it may be better to omit from the breeding programme parents that are more interactive. It has been argued, however, that this approach is less than ideal as gain is not large enough to justify the extensive progeny testing programme that is then required (Carson, 1991). Additionally, Carson (1991) found the most interactive parents can also be the best at some sites.

Regardless of whether a regionalised breeding strategy is implemented in New Zealand or not, selection of progeny testing sites is critical. Largest genetic gains will be made when testing sites display higher heritabilities (Johnson and Burdon, 1990). More importantly, Johnson and Burdon (1990) cautioned that ignoring information from a region is likely to lead to poor identification of superior clones for that region, after finding predicted volume gains dramatically reduced when information from some sites was excluded from selection analyses. The implication of this was that progeny testing in a variety of regions is essential, even if breeding programmes are not regionalised. This is a concern for the current project, given most progeny trials in the New Zealand breeding programme have been established on fertile sites close to Rotorua (Figure 1), on the premise that genotype x environment interaction has not yet been deemed to have a critical effect on selection (Shelbourne et al., 1986).

It is generally supposed that selection environments should be as similar as possible to production environments (Mulder and Bijma, 2005, Johnson and Burdon, 1990), although results of Owino and Zobel (1977b) suggest that it may be just as reliable to select in less favourable sites as it is in sites which are thought to represent a region well: it is thought unlikely that this is the case for *P. radiata* in New Zealand. There

will usually be an upper and lower limit on the number of progeny test sites: the former due to the cost of such sites, and the latter depending on the required precision for breeding values, the risk of losing a test site, and how well the site can represent the region (Carson, 1991).

1.4 Objectives

The general objective of this study was to advise on the need (or not) for partitioning breeding and deployment efforts for *P. radiata* in New Zealand, depending on the stability of genotype performance across environments.

The specific objectives were:

a) to attempt to identify those environmental variables which cause *P. radiata* genotypes to deviate significantly from predicted values; and
b) to identify sets of interacting sites and *P. radiata* genotypes that can be the focus of more comprehensive testing, thereby promoting potential partitioning of the current national breeding value prediction process into regions.

1.5 Intended impact of the research

An understanding of the interaction between genotypes and environments enables a decision to be made regarding the necessity of regionalising breeding and deployment efforts in New Zealand. As Carson (1991) explained, if the genotype x environment interaction is high, regionalising breeding efforts will result in greater genetic gains for the same expenditure. In contrast, low genotype x environment interaction implies that greatest gains could be achieved with the large selection intensities possible in a national breeding programme.

Developing a methodology for highlighting genotypes that deviate most from predicted breeding values (i.e. those that exhibit higher interaction effects) and investigating the cause of these deviations would significantly enhance breeding research. It would then be possible to make recommendations regarding the partitioning of sites into regions, or the removal of genotypes from a nation-wide breeding population.

2. CONSTRUCTING THE PROXIMITY MATRIX

2.1 Pilot analysis

In order to determine the environmental variables that are most influential on the genotype x environment interaction of *Pinus radiata* in New Zealand and to identify interacting sites and genotypes, the performances of genotypes at each site in terms of diameter at breast height were calculated. Performances were then ranked and correlations calculated between genotype rankings at all pairs of sites.

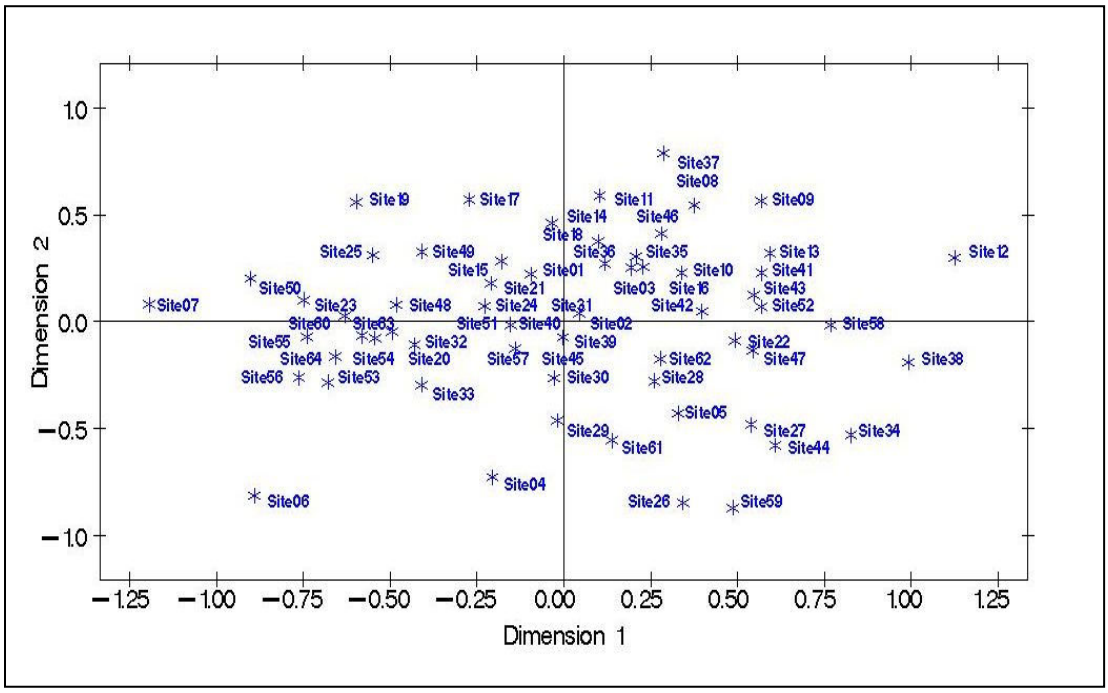


Figure 2: Multi-dimensional scaling plot of diameter at breast height family means from pilot analysis.

A pilot analysis was carried out using the MEANS procedure in SAS to calculate family means for the traits. The CORR procedure was embedded in a SAS macro to produce correlations of family means between each pair of sites. These correlations were used to produce a similarity matrix for each trait by subtracting correlations from unity. As the pilot analysis did not include experimental design features of the trial, it was foreseen to be inadequate. However, a multi-dimensional scaling plot of the diameter at breast height correlations was produced with the MDS procedure in SAS (Figure 2). The multi-dimensional scaling plot allowed a picture of the data to be

developed from an environment perspective: How does one site compare to another site in terms of the performance of families that are common to both sites? Sites that are closely plotted encourage similar performance rankings (in terms of diameter at breast height performance) of the families that are planted on them. Conversely, sites plotted further apart contain family rankings that differ from each other.

The R statistical software was used to produce a level plot of female clone performance across sites in terms of diameter at breast height (Figure 3). The intention of this plot was to provide a picture of the data from the genotype perspective: How does one particular family perform on one site compared to its performance at all other sites where it is grown?

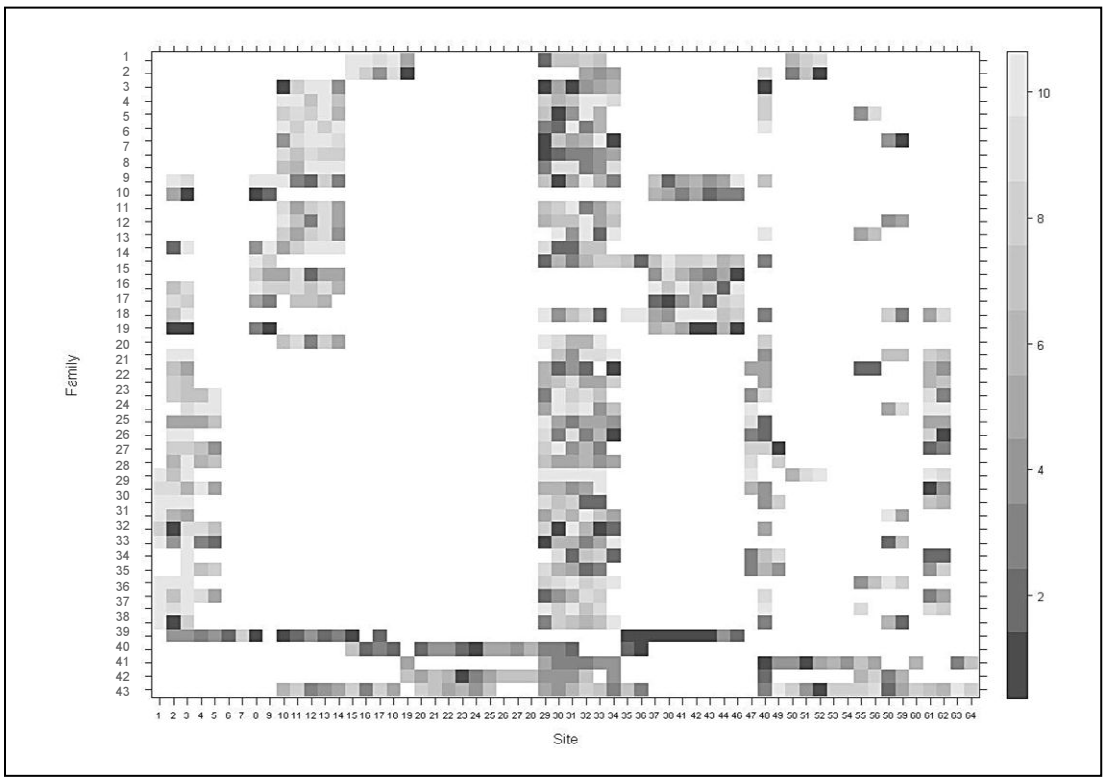


Figure 3: Level plot of family performance (diameter at breast height) across sites from pilot analysis.

Each female clone was assigned to a decile group at each site at which its offspring were present. A clone's decile grouping across all sites is represented as a row of colour-coded cells. A clone with predominantly darker cells is considered to be a poor performer across all sites, whereas a clone with predominantly brighter cells is considered to be a good performer across all sites. Clones with mixed performances

are expressing some form of genotype x environment interaction. Due to space constraints, the level plot was limited to the most-represented female clones, i.e. those female clones whose offspring appear at more than ten sites.

The multi-dimensional scaling and level plots were used to guide the development of the final methodology for construction and analysis of the proximity matrix. Construction of a more formal proximity matrix was expected to enable the specific objectives to be achieved in conjunction with analyses using threshold graphs and multiple regression on resemblance matrices that would stem from the use of the matrix.

2.2 Calculating correlations between performance rankings

A formal analysis was conducted using ASReml-R, a statistical software that easily allowed utilisation of breeding values and incorporation of experimental design features in the calculation of correlations between sites. Although Burdon's Type-B genetic correlations have been widely used for estimating between-site genetic correlations, this method has been superseded the multivariate evaluation used here.

The RPBC data was imported into ASReml-R where diameter at breast height measurements were analysed two sites at a time. During each loop, a univariate analysis (incomplete block design) was conducted on the data from each of the two sites using residual maximum likelihood to produce variance components. For the univariate analysis, the model fitted to each site depended on both mating design and experimental design using the following model:

$$y = Xm + Z_S s + Z_P p + Z_T a + Z_F f + \varepsilon \qquad [2]$$

where y is the vector of observed traits for the i^{th} site, m is the vector of fixed effects including the overall trait mean and the replicate effects, s is the vector of random set effects nested within replicate effects, p is the vector of random plot effects nested within replicate effects, a is the vector of random additive genetic effects, f is the vector of random family effects and ε is the vector of random residuals. X, Z_S, Z_P, Z_T,

and Z_F are incidence matrices relating *m*, *s*, *p*, *a,* and *f* to *y*. The expected value and dispersion matrices assuming a multivariate normal distribution (MND) are:

$$\begin{bmatrix} y \\ s \\ p \\ a \\ f \\ \varepsilon \end{bmatrix} \sim MND\left(\begin{bmatrix} Xm \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} V & Z_S S & Z_P P & Z_T G & Z_F F & R \\ S\ Z'_S & S & 0 & 0 & 0 & 0 \\ P\ Z'_P & 0 & P & 0 & 0 & 0 \\ G\ Z'_T & 0 & 0 & G & 0 & 0 \\ F\ Z'_F & 0 & 0 & 0 & F & 0 \\ R & 0 & 0 & 0 & 0 & R \end{bmatrix}\right) \quad [3]$$

where $S = \sigma_s^2 I$, $P = \sigma_p^2 I$, $G = \sigma_a^2 A$, and $R = \sigma_\varepsilon^2 I$ are the set, plot, additive genetic, and residual covariance matrices, respectively, *A* is the numerator relationship matrix, and *0* is a null matrix (with all elements equal to 0). The phenotypic covariance matrix is:

$$V = Z_S s\ Z'_S + Z_P p\ Z'_P + Z_T a\ Z'_T + Z_F f\ Z'_F + \varepsilon \quad [4]$$

Variance components from the two univariate analyses were then iteratively extracted from the variance components' tables. The variance components were used to run a bivariate analysis between all pairs of sites, and the genetic correlations between all pairs of sites were extracted. For the bivariate analysis the following model was used:

$$y = Xm + Z_S s + Z_P p + Z_T a + Z_F f + \varepsilon \quad [5]$$

where *y* is the vector of observed traits for a pair of sites, *m* is the vector of fixed effects including the overall trait mean, the *i* site effects, and the *j* replicate effects nested within sites, *s* is the vector of random set effects nested within replicates nested within sites, *p* is the vector of random plot effects nested within replicates nested within sites, *a* is the vector of random additive genetic effects nested within sites, and ε is the vector of random residuals. *X*, Z_S, Z_P, and Z_T are incidence matrices relating *m*, *s*, *p*, and *a* to *y*. The expected value and dispersion matrices assuming a multivariate normal distribution (MND) are:

$$\begin{bmatrix} y \\ s \\ p \\ a \\ f \\ \varepsilon \end{bmatrix} \sim MND\left(\begin{bmatrix} Xm \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} V & Z_S S & Z_P P & Z_T G & Z_F F & R \\ S\ Z'_S & S & 0 & 0 & 0 & 0 \\ P\ Z'_P & 0 & P & 0 & 0 & 0 \\ G\ Z'_T & 0 & 0 & G & 0 & 0 \\ F\ Z'_F & 0 & 0 & 0 & F & 0 \\ R & 0 & 0 & 0 & 0 & R \end{bmatrix}\right) \quad [6]$$

where $S = \sum_{\oplus} \sigma^2_{s_i} I_i$, $P = \sum_{\oplus} \sigma^2_{p_i} I_i$, $G = A \otimes G_0$ and $R = \sum_{\oplus} \sigma^2_{\varepsilon_i} I_i$ are the set, plot, additive genetic, and residual covariance matrices, respectively, A is the numerator relationship matrix, 0 is a null matrix (with all elements equal to 0), $\Sigma_{\oplus}$ denotes a direct sum, $\otimes$ represents a direct product operation (Searle, 1982), and:

$$G_0 = \begin{bmatrix} \sigma^2_{a_1} & \sigma_{a_{12}} \\ \sigma_{a_{12}} & \sigma^2_{a_2} \end{bmatrix} \quad [7]$$

2.3 Constructing and visualising the proximity matrix

Once the genetic correlations had been extracted for all pairs of sites, the IML procedure in SAS was used to populate a proximity matrix showing the "distance" between any two sites in terms of the genetic correlation for diameter at breast height. Distances to be used in the matrix were derived by subtracting the correlations from unity.

Of the 2850 cells ((n x (n-1))/2) in each triangular portion of the proximity matrix, just 219 were populated, for which the bivariate analyses of 70 pairings did not converge in ASReml-R (Appendix B). Thus, just over 5 % (148 out of 2850) of the site pairings were sufficiently connected to enable derivation of a genetic correlation. However, of the 76 sites for which there was data available, all but one site (FR216_1) was connected to at least one other site.

A frequency chart of the genetic correlations is shown in Figure 4. The distribution is left-skewed, meaning that there are very few negative correlations. This indicates a low occurrence of ranking reversals. About one-third of the correlations are greater than 0.6, suggesting minimal genotype x environment interaction. More importantly, nearly two-thirds (95 out of 148) of the correlations lie between -1.0 and 0.6, highlighting the presence of strong genotype x environment interactions. This result supports the type-B genetic correlations for diameter at breast height found by Jayawickrama (2001) for data from New Zealand and New South Wales.

Having quantified the genotype x environment interaction using these measures of proximity, multi-dimensional scaling was used as in the pilot analysis to gain a visual representation of the similarity of sites.

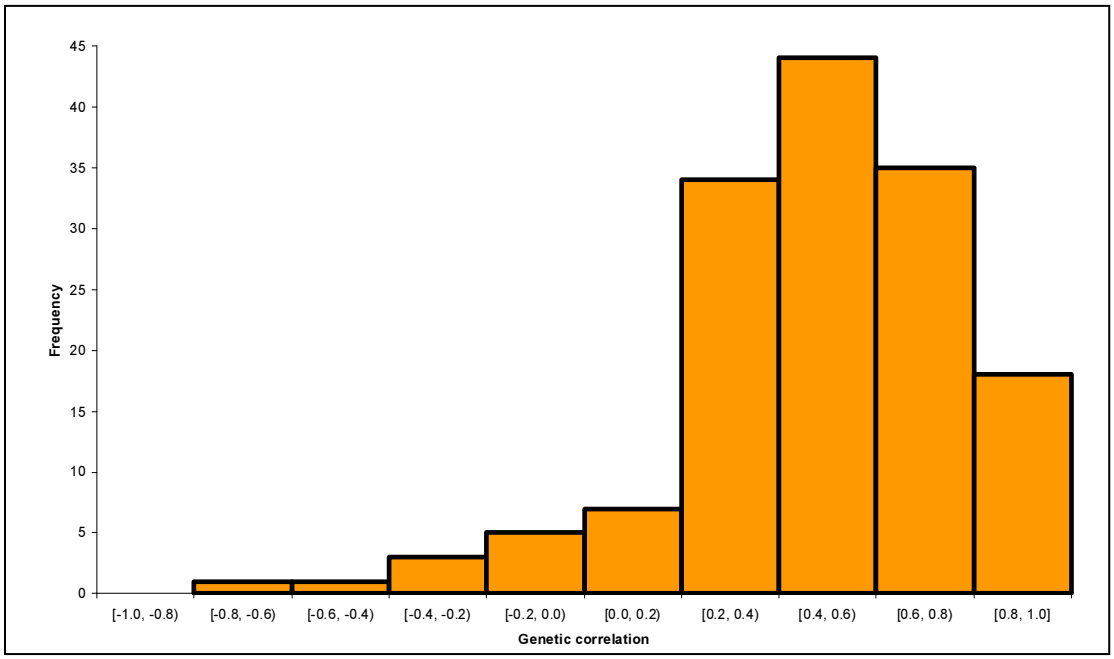


Figure 4 – Frequency chart of genetic correlations for 148 site pairs.

The multi-dimensional scaling plot shows a large number of sites clumped around the origin, with some sites scattered around the outside of the plot (Figure 5). An explanation for the placement of sites on the plot was not apparent, neither geographically nor climatically. However, it is possible that the clustering of points around the origin is due to there being insufficient between-site connectivity to enable a more decisive separation of sites.

A level plot was used to view genotype performance across sites, for families where the parent clones are in the current production population (Figure 6). However, in this case, breeding values were extracted from ASReml-R's *.sln* output files. The level plot shows a number of families performing consistently across sites. A number of families plotted (268538, 288302, 886947, 886944, 886925, 886917, 886881, 886945, 880729, and 288402) performed consistently poorly on the sites they were present, showing on the plot as having a majority of pink-coloured squares. In contrast, there were several families (875012, 887734, 883005, 883058, 885212, 850533, 885458,

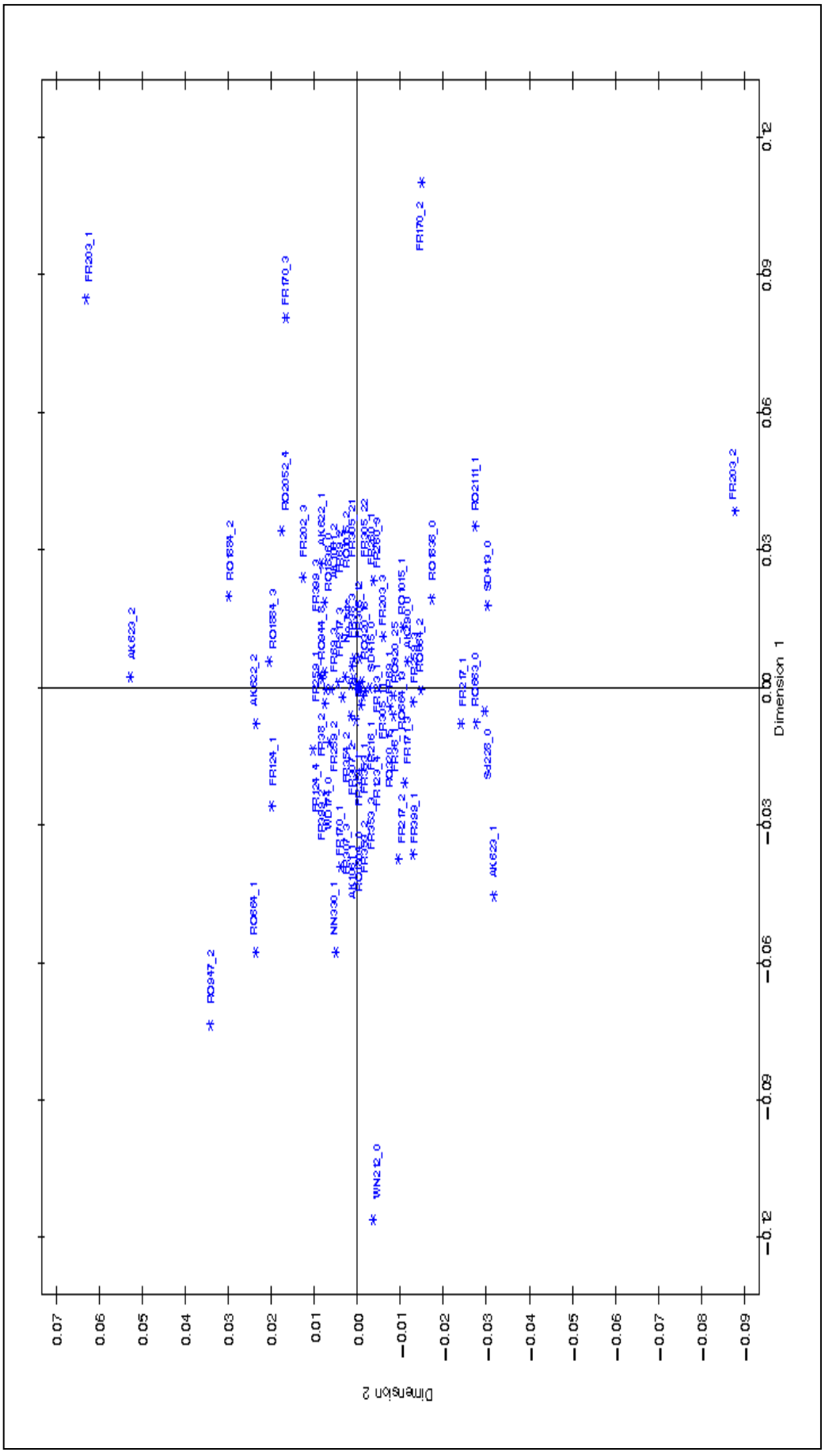


Figure 5: Multi-dimensional scaling plot of diameter at breast height.

Figure 6: Level plot of family performance (diameter at breast height) across sites using parents in production population (Proseed).

885280, and 268609) that performed consistently well across sites, having mostly green-coloured squares.

Just under two-thirds of the families plotted on the level plot displayed large rank changes across sites. That is, at some sites the families were ranked in higher decile groupings, whereas at other sites those same families were ranked in lower decile groupings. These families are expressing the effect of genotype x environment interactions, reinforcing the result found from Figure 4. For example, family 880606 ranks in the top decile of performers (green cells) at sites FR202_3, FR217_2, WN212_0, RO1836_0, and RO1884_2. However, it is in the lowest decile (pink cells) at sites FR307_2, FR217_3, FR217_1, and FR171_3.

3. ISOLATING THE ENVIRONMENTAL VARIABLES THAT ARE DRIVING GENOTYPE X ENVIRONMENT INTERACTIONS

3.1 Extraction of environmental variables

Environmental data were extracted from the National Climate Database (CliDB) operated by New Zealand's National Institute of Water and Atmospheric Research (NIWA) using the CliFlo web service[3]. NIWA store data from numerous weather stations around New Zealand and the Pacific. Hawth's geospatial analysis tools in ArcMAP® 9.2 (ESRI, 2006) were used to select weather stations situated closest to each trial's GPS location. In some cases, weather stations that were further from the trial coordinates than the closest station were selected, in order to ensure that sufficient data for climatic variables were available. The average straight line distance a weather station was from a trial was 27 km, with a maximum distance of 103 km.

As many variables as were available were extracted from CliDB for the period between and including the year of establishment of trial and the time of measurement of the trial. It was decided to restrict variables used in the analysis to those that were available for most sites. This was just a small collection of variables, including: mean monthly rainfall (mm), mean air temperature (°C), mean daily maximum air temperature (°C), mean daily minimum air temperature (°C), extreme maximum air temperature (°C), extreme minimum air temperature (°C), mean vapour pressure (hPa), and maximum 24-hour rainfall (mm). For many trials, a combination of 2 or 3 weather stations were used to ensure climatic data were available for the desired time period. In addition, altitude was calculated by intersecting the trial GPS coordinates with the underlying raster from Landcare Research's Digital Elevation Model.

3.2 Multiple regression on proximity matrix

Dissimilarity matrices were constructed for each environmental variable by calculating the absolute distance between two sites as the value at the first site subtracted from the value at the second site. These (independent or explanatory)

[3] http://cliflo.niwa.co.nz/ accessed between 6th August 2008 and 5th September 2008

environmental matrices were then compared with the (dependent or response) proximity matrix of genetic correlations for diameter at breast using a procedure known as multiple regression on resemblance matrices or MRM (Legendre et al., 1994, Lichstein, 2007, Smouse et al., 1986).

MRM has evolved from the work of Mantel (1967), who was investigating time-space clustering of leukemia. Mantel (1967) generated statistical significance levels for the association between time and space distances between pairs formed from observed cases of disease. Since then, this general procedure for matrix comparison has been applied and extended by a number of authors in a number of fields of research (Douglas and Endler, 1982, Dow and Cheverud, 1985, Hubert and Golledge, 1981, Legendre et al., 1994, Lichstein, 2007, Schnell et al., 1985, Smouse et al., 1986).

Compared to traditional Mantel analysis, MRM offers the opportunity of separating environmental variables into individual distance matrices to allow inferences to be made about these variables without fear of dilution by unimportant variables (Lichstein, 2007). Here, it was used to assess the impact of various environmental variables on the performance of genotypes as measured by phenotypic growth responses. In addition, calculations for fitting an MRM are the same as those for a multiple regression with standard datasets. However, due to dependence issues in a distance matrix, significance of results for MRM is usually tested through permutation rather than using Fisher's Z-transformation (Dow and Cheverud, 1985).

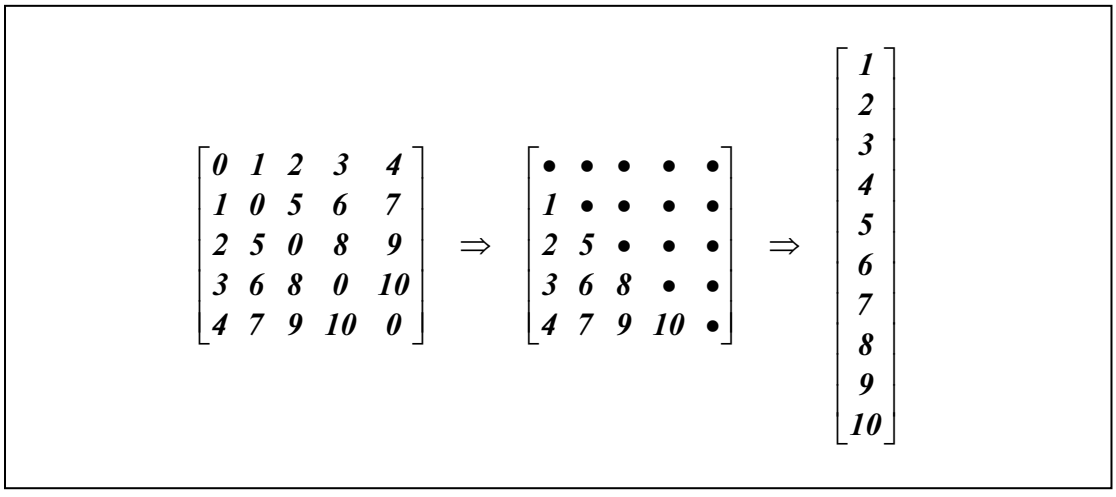


Figure 7: Schematic of "unfolding" procedure (for a 5 x 5 matrix) used in MRM analysis.

Each matrix was symmetric, so that the upper right triangle and the lower left triangle of the matrix were reflections of each other (Figure 7). As the entries on the main diagonal represented the distance between a site and itself, they were all zero. Therefore, one of the triangular portions of each matrix was regarded as redundant and the main diagonal (containing self-distances) was discarded, leaving $n(n\text{-}1)/2$ distances. The remaining distances in all matrices were then unfolded in the same sequence to form vectors of distances. Each vector of distances was then regressed against the explanatory distance vectors using the GLM procedure in SAS resulting in a series of t-statistics.

Tests of significance of these t-statistics could not be performed with the traditional parametric methodology, due to the lack of independence between observations inherent in a dissimilarity matrix (Legendre et al., 1994). Therefore, the rows and columns of the genetic correlation proximity matrix were randomly permuted 2500 times. At the end of each iteration, t-statistics were calculated, randomly assigned as either positive or negative, and used to develop null distributions for the test statistic of each explanatory variable (Appendix C).

The CAPABILTY procedure in SAS was then used to test if these null distributions were normal using three empirical distribution function tests of normality: Kolmorgorov-Smirnov, Anderson-Darling, and Cramer-von Mises. The results of these tests are summarized in Table 1.

There was no evidence that any of the null distributions were centred on a value different from zero, as seen by the non-significant results in the tests for location. More importantly, all null distributions were shown to be normal by all three tests of normality with the exception of the Kolmogorov-Smirnov test for maximum temperature.

This finding allowed the results of the MRM analysis to be tested against a normal distribution. That is, a *t*-statistic greater than 1.96 could be considered significant with 95 % confidence, and a *t*-statistic greater than 1.645 could be considered significant with 90 % confidence.

Table 1 – Results of normality tests for generated null distributions of critical *t*-values.

	Altitude	Total rainfall	24-hr rainfall	Vapour Pressure	Mean temperature	Extreme maximum temperature	Extreme minimum temperature	Maximum temperature	Minimum temperature
Moments									
Mean	-0.014	0.042	0.015	-0.010	-0.004	-0.007	-0.004	-0.018	-0.009
Std Deviation	1.041	1.001	1.009	1.028	0.995	1.021	0.989	0.993	0.994
Skewness	-0.067	0.046	-0.020	-0.035	-0.063	0.023	0.024	-0.018	0.001
Variance	1.083	1.003	1.018	1.056	0.990	1.043	0.978	0.985	0.989
Kurtosis	0.090	0.042	-0.008	-0.053	0.428	0.110	-0.100	0.343	0.238
Tests for location Mu0=0									
Student's *t* (*t*)	-0.680	2.086	0.733	-0.471	-0.186	-0.343	-0.212	-0.914	-0.460
Pr ≥ |*t*|	0.496	0.037	0.464	0.638	0.853	0.731	0.832	0.361	0.645
Sign (M)	3	28	28	9	13	7	-8	-32	-30
Pr ≥ |M|	0.920	0.271	0.271	0.734	0.617	0.795	0.764	0.208	0.238
Signed Rank (S)	-15475	68087	29556	-10616	-986	-10610	-8901	-25238	-20770
Pr ≥ |S|	0.668	0.059	0.413	0.769	0.978	0.769	0.805	0.485	0.565
Tests for normality									
Kolmogorov-Smirnov (D)	0.013	0.012	0.015	0.016	0.012	0.011	0.009	0.019	0.013
Pr > D	>0.150	>0.150	>0.150	0.139	>0.150	>0.150	>0.150	0.029 *	>0.150
Cramer-von Mises (W^2)	0.064	0.038	0.074	0.068	0.064	0.041	0.019	0.115	0.093
Pr > W^2	>0.250	>0.250	>0.250	>0.250	>0.250	>0.250	>0.250	0.075	0.142
Anderson-Darling (A^2)	0.435	0.234	0.488	0.379	0.486	0.273	0.143	0.705	0.564
Pr > A^2	>0.250	>0.250	0.23	>0.250	0.232	>0.250	>0.250	0.069	0.148

* Significant at α = 0.05

3.3 Results of multiple regression on proximity matrix

Correlation-based distances from the 148 site pairs that successfully converged, were plotted against the nine environmental variables (Figure 8 through Figure 16). All graphs display a 'frontier'-type pattern, with a large clustering of points in the lower left quadrant (near the origin), some points in the upper left and lower right quadrants, and virtually no points in the upper right quadrant.

This pattern should not be unexpected: due to most of the sites in this study being based in the Central North Island, climatic (or macro-environmental) variables for these sites should not be too dissimilar. This explains the tendency for environmental distances for site pairs to be grouped closer to the origin.

In addition, correlation-based distances greater than one imply that performances of genotypes on one site are opposite to performances of those same genotypes on another site. While changes in rank have been observed, complete reversals of rank are thought to be uncommon (Matheson and Raymond, 1984). It is, therefore, reasonable that only a few site pairs appear above unity in any of the graphs.

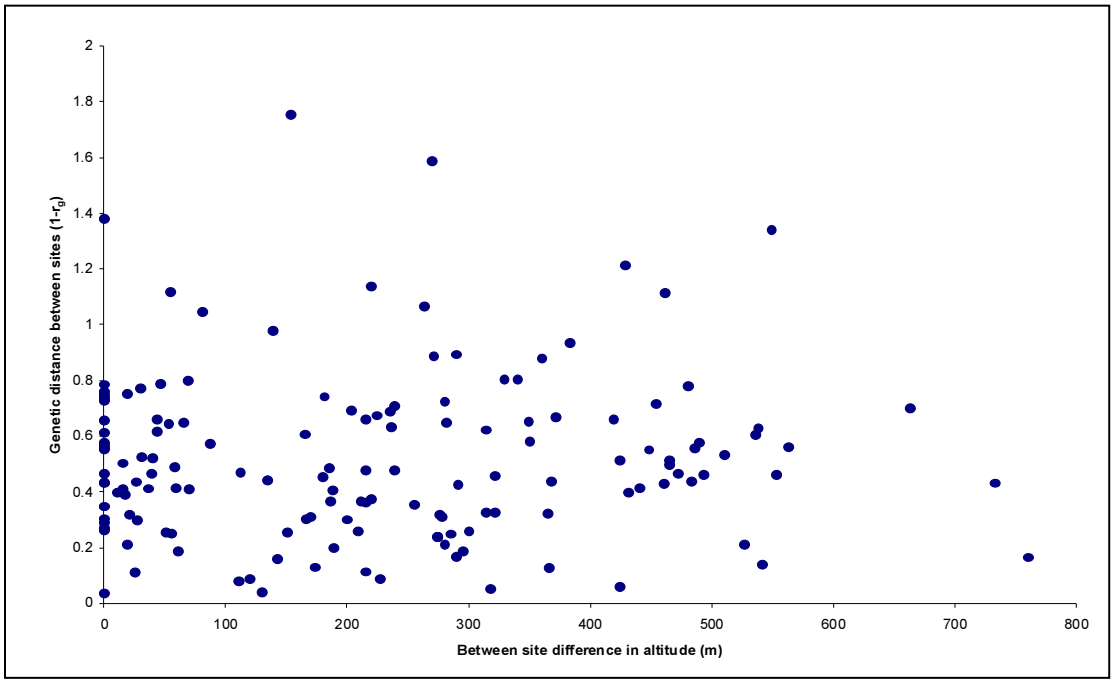


Figure 8 – Relationship between correlation-based distance and between-site difference in altitude for 148 site pairs.

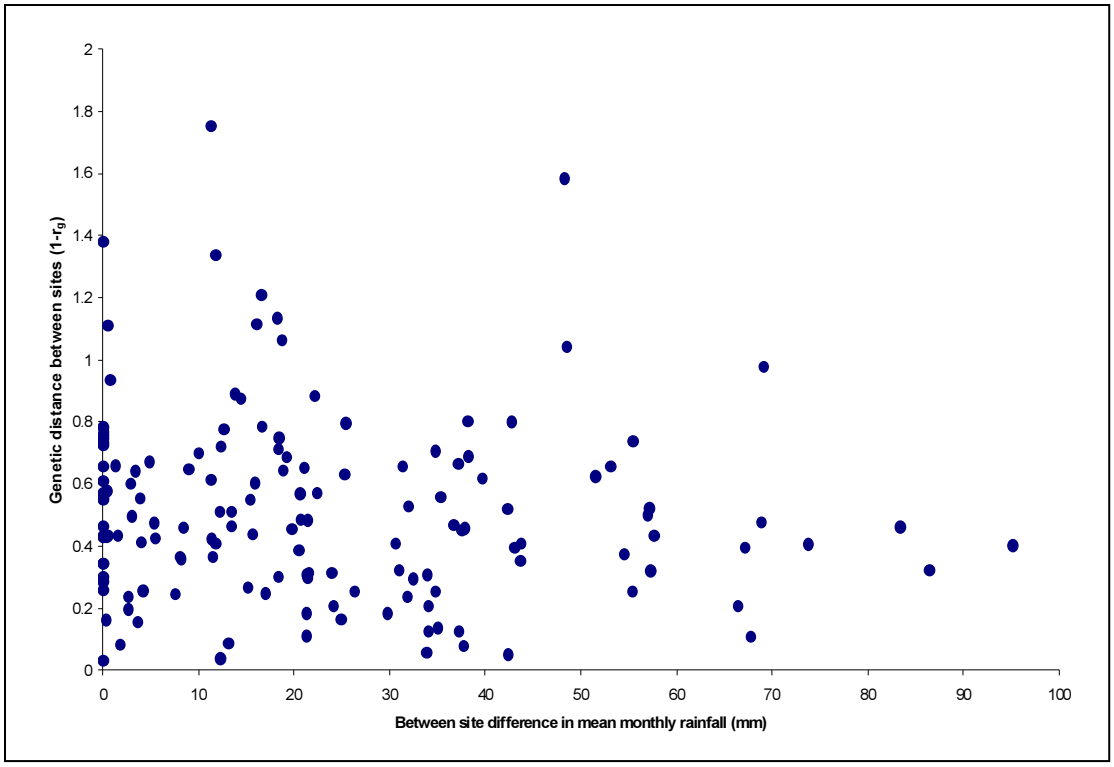


Figure 9 – Relationship between correlation-based distance and between-site difference in mean monthly rainfall for 148 site pairs.

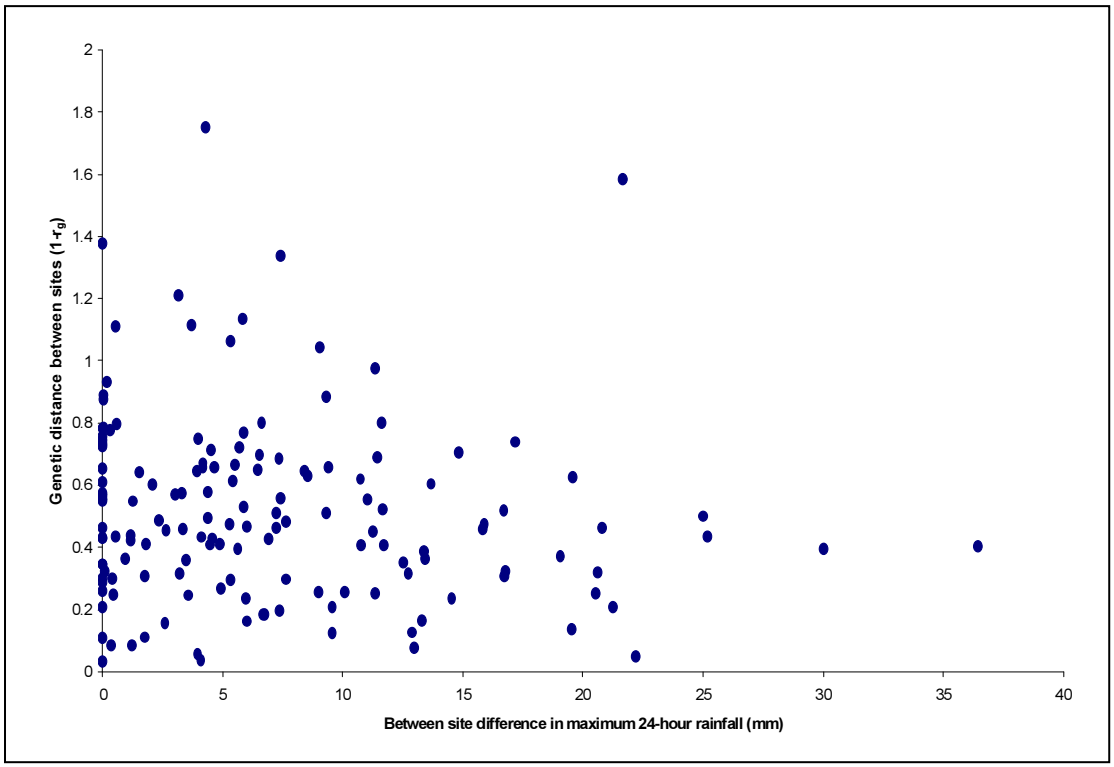


Figure 10 – Relationship between correlation-based distance and between-site difference in maximum 24-hour rainfall for 148 site pairs.

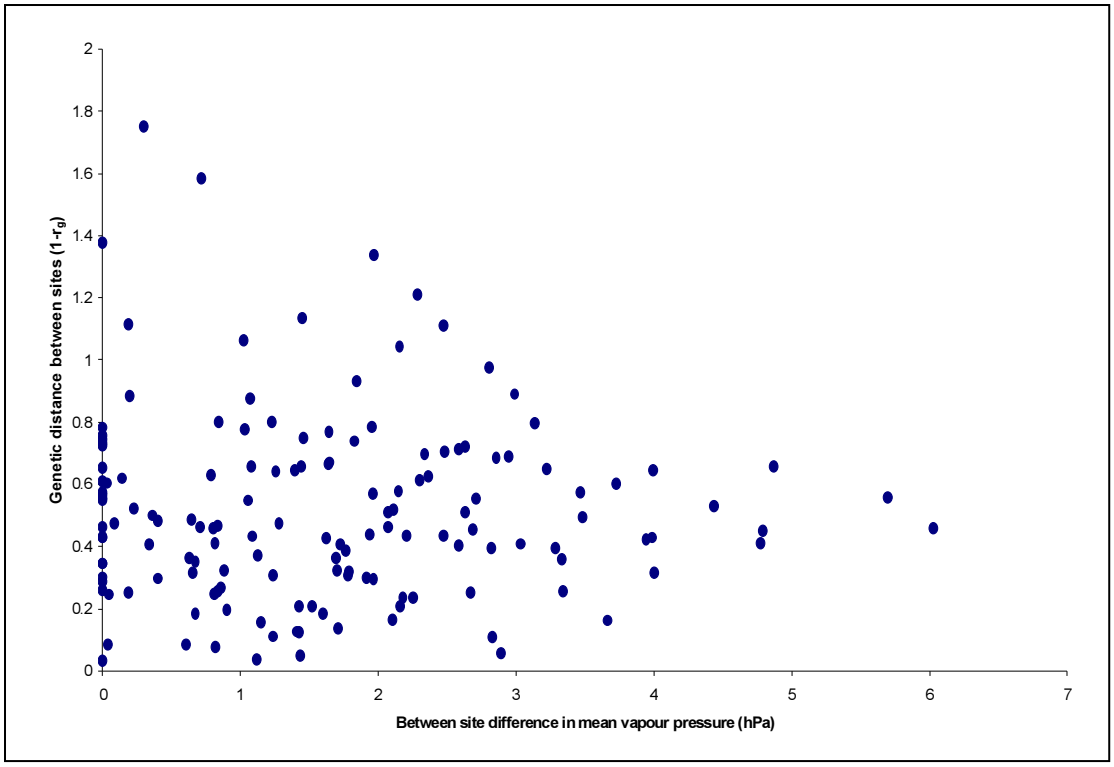


Figure 11 – Relationship between correlation-based distance and between-site difference in mean vapour pressure for 148 site pairs.

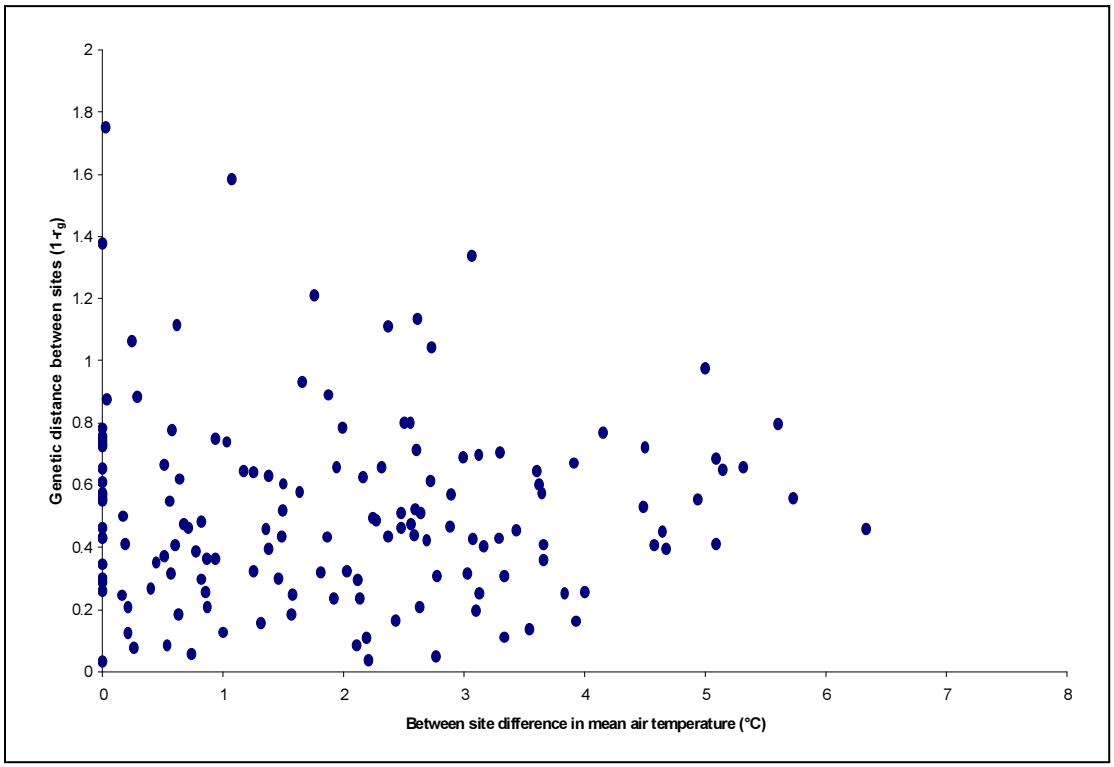


Figure 12 – Relationship between correlation-based distance and between-site difference in mean air temperature for 148 site pairs.

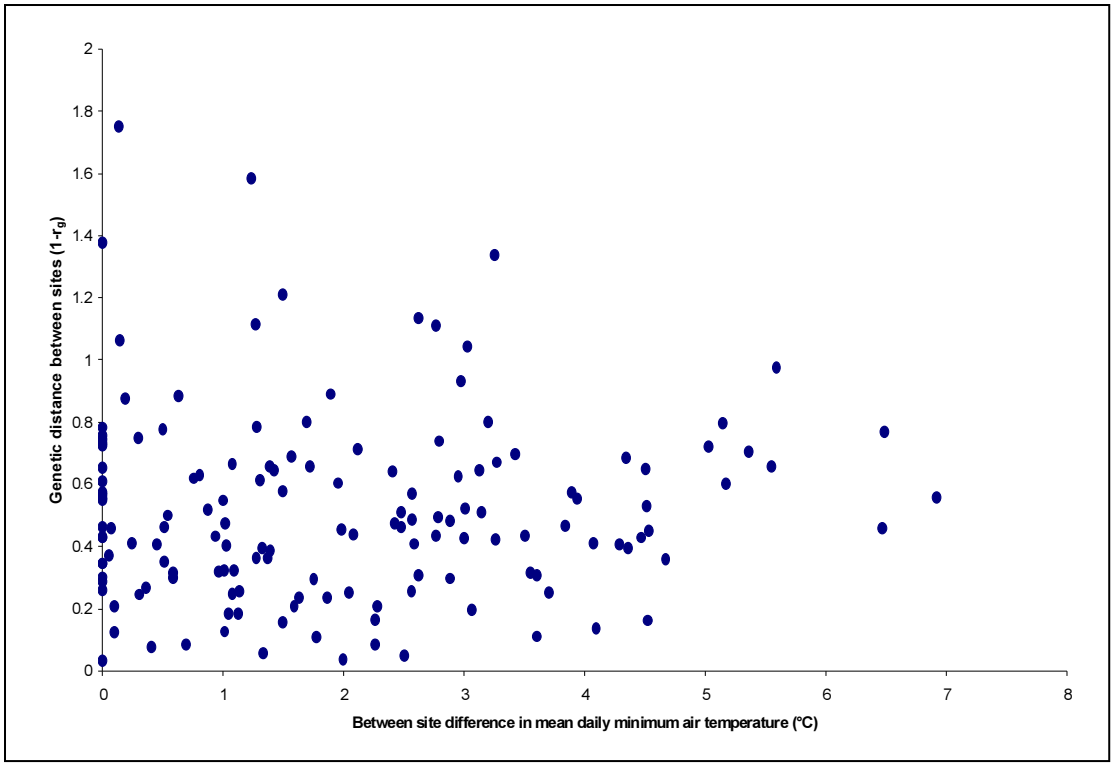


Figure 13 – Relationship between correlation-based distance and between-site difference in mean daily minimum air temperature for 148 site pairs.

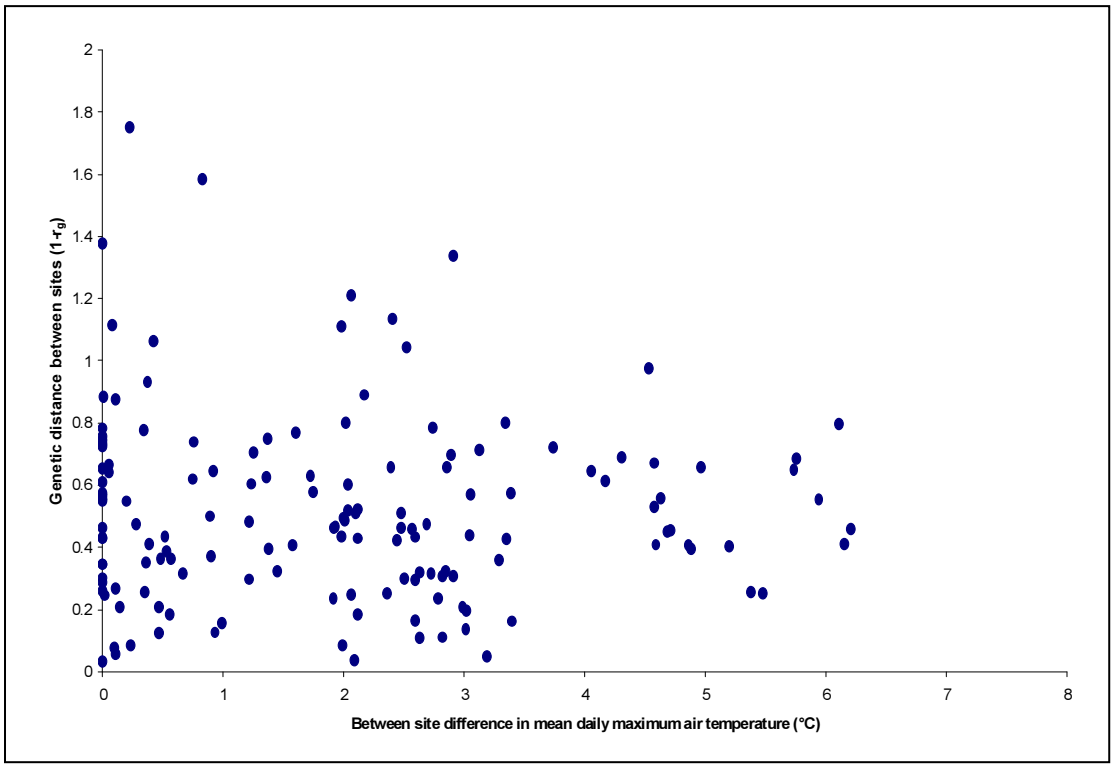


Figure 14 – Relationship between correlation-based distance and between-site difference in mean daily maximum air temperature for 148 site pairs.

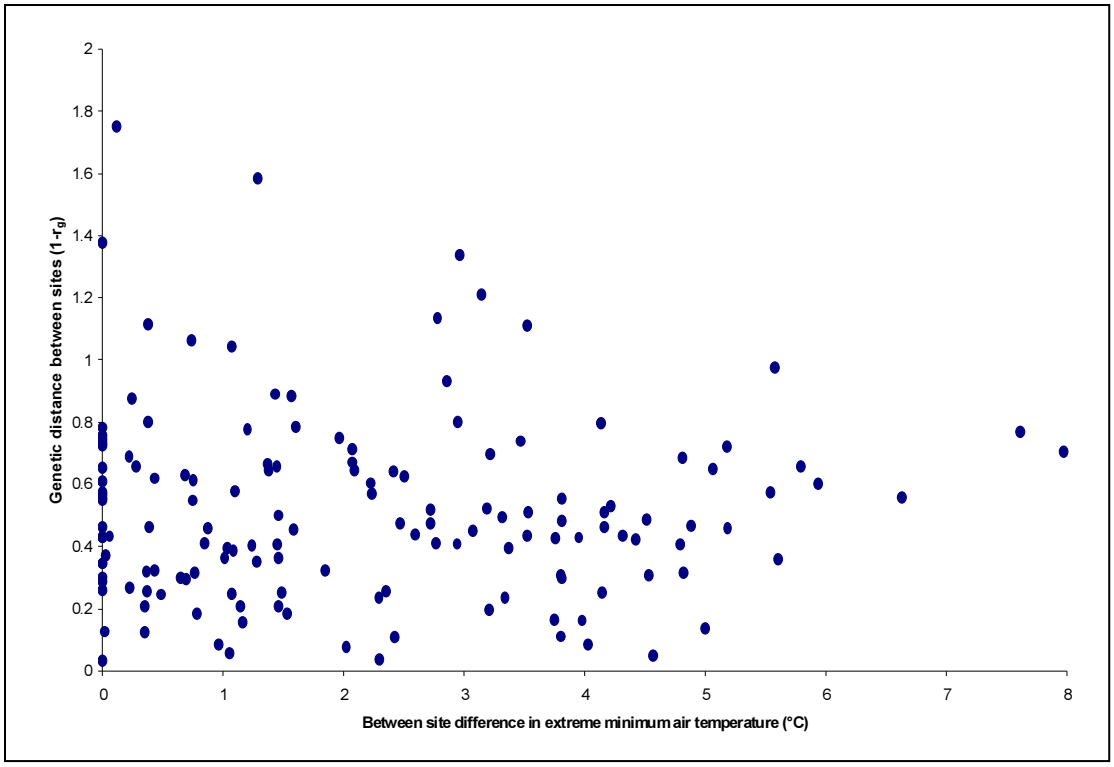


Figure 15 – Relationship between correlation-based distance and between-site difference in extreme minimum air temperature for 148 site pairs.

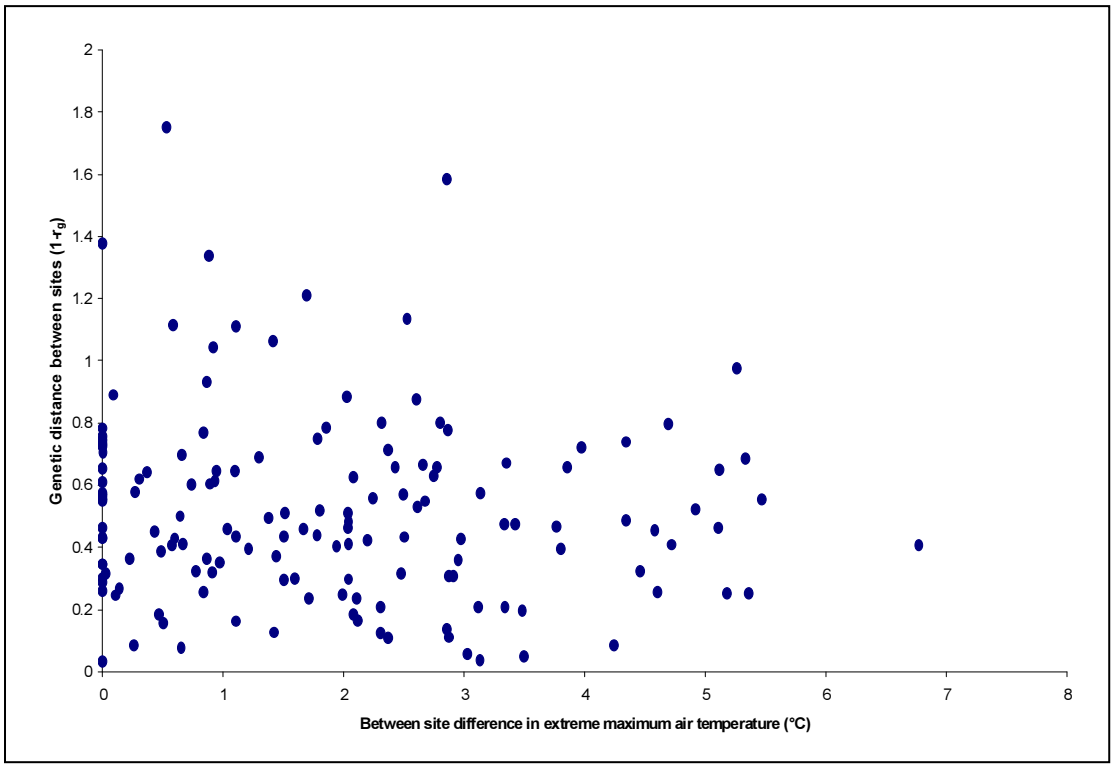


Figure 16 – Relationship between correlation-based distance and between-site difference in extreme maximum air temperature for 148 site pairs.

The regression analysis showed genetic correlations for diameter at breast height to be significantly affected by extreme maximum temperatures, with a *t*-statistic of 2.03 (Table 2 and Appendix D). Despite not assigning a probability to this *t*-statistic, it can be stated that there is at least 95 % confidence in this result, as the *t*-statistic exceeds the critical *t*-value of 1.96 (for $\alpha = 0.05$).

Table 2 – Result of regression procedure for diameter at breast height genetic correlation vector regressed against environmental variable vectors.

Source	*DF*	*Sum of Squares*	*Mean Square*	*F-value*
Model	1	0.0771	0.0771	4.11
Error	2276	42.7030	0.0188	
Corrected Total	2277	42.7801		

R-Square	*Coeff Var*	*Root MSE*	*SiteDist Mean*
0.0018	481.9978	0.1370	0.0284

Source	*DF*	*Type III SS*	*Mean Square*	*F-value*
ExMaxTemp	1	0.0771	0.0771	4.11

Parameter	*Estimate*	*Standard Error*	*t-value*
Intercept	0.0206	0.0048	4.26
ExMaxTmp	0.0042	0.0021	2.03

Altitude was the last of the environmental variables to be excluded from the regression analysis (Appendix D). However, when tested on its own, there is at least 90 % confidence that altitude affects the genetic correlations between sites (Table 3).

Table 3 – Result of regression procedure for diameter at breast height genetic correlation vector regressed against altitude vectors.

Source	*DF*	*Sum of Squares*	*Mean Square*	*F-value*
Model	1	0.0689	0.0689	3.67
Error	2276	42.7113	0.0188	
Corrected Total	2277	42.7801		

R-Square	*Coeff Var*	*Root MSE*	*SiteDist Mean*
0.0016	482.0443	0.1370	0.0284

Source	*DF*	*Type III SS*	*Mean Square*	*F-value*
Altitude	1	0.0689	0.0689	3.67

Parameter	*Estimate*	*Standard Error*	*t-value*
Intercept	0.0207	0.0049	4.20
Altitude	0.0000	0.0000	1.92

4. IDENTIFYING INTERACTING SITES AND GENOTYPES

4.1 Clustering using proximity thresholds

Due to the sparseness of the proximity matrix, many more common clustering techniques were unable to be applied to the RPBC data. However, use of some basic definitions from Graph Theory allowed the development of an algorithm for grouping sites.

An adjacency matrix is defined as the n x n matrix in which the entry in row i and column j is the number of edges joining the vertices i and j (Aldous and Wilson, 2000). In the case of the RPBC data, the vertices i and j are represented in the adjacency matrix as sites i and j, and the edge joining the vertices can be thought of as a "genetic connection" between sites i and j. Therefore, a cell in the proximity matrix containing a correlation-based distance was thought of as displaying a connection between two sites.

For the purposes of this study, the existence of an edge, or connection, between two sites was thought more important than the exact "length" of that connection *per se*. Therefore, the proximity matrix was converted to a form of adjacency matrix by setting an initial threshold of 0.02 (for genetic correlation-based distances) and replacing threshold-bound entries with unity, indicating an incidence or strong connection between the two sites. Missing entries and threshold-exceeding entries were replaced with zero to represent no or low incidence (or connection) between those two sites.

It is also true that the number of walks of length k from vertex i to vertex j is equal to the entry in row i and column j of the k^{th} power of the adjacency matrix (Aldous and Wilson, 2000). For this reason, the adjacency matrix was then multiplied by the n^{th} (i.e. 76^{th}) power, replacing all positive entries with unity at each matrix-multiplication, to give blocks of connectivity. From the resulting matrix, clusters were identified as those row (or column, due to symmetry) numbers in each column (or row) where the matrix entry was unity.

The threshold was then slackened by 0.01 and the process repeated. Each change in threshold presented an expanded set of clusters. That is, more sites were included in existing clusters, new sites formed a cluster of their own, and/or two or more clusters merged into one cluster.

4.2 Results of clustering using proximity thresholds

The full results of the clustering analysis are shown in Appendix E. No sites are bound by the initial threshold of 0.02. Similarly, a threshold of 0.03 does not bound any sites. At a threshold of 0.04, just two sites 58 (FR307_2) and 59 (FR307_3) are bound by the threshold. However, these two sites do not form a cluster. They are simply connected with themselves as single-site clusters. The threshold was progressively slackened, including more and more sites as single-site clusters.

At a threshold of 0.09, the first multi-site clusters began to form. Site 10 (AK622_1) was found to be connected with site 58 (FR307_2), site 32 (FR203_1) connected with site 59 (FR307_3), and site 34 (FR203_3) connected with 47 (RO2052_4). More and larger clusters continued to be formed until the threshold was relaxed to 0.72 (Figure 17). From this point, there was no further change to the clusters.

In general, amalgamation algorithms continue clustering until all objects belong to a single class (Gordon, 1999). It is then up to the user to determine a “stopping rule” to decide how many clusters there should be (Milligan and Cooper, 1985). However, due to the lack of connectivity on the proximity matrix, the threshold methodology was unable to establish even the most distant connection between all sites. Therefore, it was thought that a stopping rule was not necessary in this analysis. Instead, the clustering was allowed to continue until all sites were included in a cluster and the clusters stabilised (i.e. at a threshold value of 0.72).

The final make-up of the clusters reflected the trial series from which the cluster members were a part of. For example, sites FR38_1, FR38_2, and FR38_3 formed a three-site cluster. These sites were all planted in 1988 as part of the same trial series, and presumably with the same (or very similar) genetic material.

Site number	Trial id	Wood region	Northing	Easting	Altitude	Year of Establishment	Year of Measurement	Age at Measurement
1	RO944_8	Central North Island	6318820	2832690	415	1968	1978	10
2	WN212_0	Hawkes Bay	6158726	2796271	780	1969	1979	10
3	RO947_2	Central North Island	6262553	2781916	466	1969	1979	10
4	AK290_0	Northland	6504000	2629358	47	1971	1984	13
5	Sd228_0	Otago/Southland	5453769	2286340	227	1971	1984	13
6	RO1015_1	Central North Island	6258790	2781600	499	1972	1981	9
7	RO320_16	Central North Island	6298060	2816510	554	1972	1982	10
8	RO1015_2	Central North Island	6258116	2781760	518	1972	1982	10
9	RO320_15	Central North Island	6298370	2815720	548	1972	1982	10
10	AK622_1	Northland	6523042	2618011	27	1975	1983	8
11	AK623_1	Auckland	6431268	2710664	80	1975	1983	8
12	NN330_1	Nelson/Marlborough	5947614	2495849	590	1975	1983	8
13	RO664_1	Central North Island	6281630	2801340	563	1975	1984	9
14	SD415_0	Otago/Southland	5467683	2271120	520	1975	1983	8
15	RO1836_0	Central North Island	6352370	2821040	280	1981	1989	8
16	RO1884_3	Central North Island	6297860	2816460	554	1983	1990	7
17	RO1838_0	Central North Island	6248133	2781409	580	1981	1985	4
18	RO1884_2	Central North Island	6307253	2773875	259	1983	1989	6
19	FR171_3	Central North Island	6334665	2827645	94	1992	2000	8
20	AK1061_1	Northland	6536939	2611689	100	1987	1995	8
21	AK1061_2	Northland	6603206	2587078	100	1987	1993	6
22	RO2111_1	Central North Island	6281100	2801220	565	1987	1995	8
23	FR38_1	Northland	6537300	2609610	81	1988	1996	8
24	FR38_2	Central North Island	6321505	2761135	372	1988	1996	8
25	FR38_3	Central North Island	6362670	2808150	98	1988	1995	7
26	FR69_1	Northland	6537300	2609610	81	1989	1998	9
27	FR69_2	Central North Island	6322545	2759615	357	1989	1997	8
28	FR69_3	Central North Island	6312985	2834650	296	1989	1996	7
29	FR170_1	Northland	6492870	2635580	20	1992	2001	9
30	FR170_2	Central North Island	6327278	2772802	509	1992	2000	8
31	FR170_3	Central North Island	6312709	2835293	300	1992	2000	8
32	FR203_1	Hawkes Bay	6212530	2829150	451	1993	2001	8
33	FR203_2	Australia	.	.		1993	2000	7
34	FR203_3	Central North Island	6314034	2830097	420	1993	2001	8
35	RO1804_0	Central North Island	6344950	2837800	20	1980	1987	7
36	RO664_13	Central North Island	6281680	2801266	561	1980	1987	7
37	AK622_2	Auckland	.	.	.	1975	1983	8
38	AK623_2	Auckland	6431390	2710715	100	1975	1983	8
39	*No site*	.	.	.	.	.	.	.
40	*No site*	.	.	.	.	.	.	.
41	RO320_25	Central North Island	6298060	2816510	554	1975	1984	9
42	RO663_0	East Coast	6370745	2973785	370	1975	1983	8
43	RO664_2	Central North Island	6281680	2801260	561	1975	1984	9
44	SD413_0	Otago/Southland	5461200	2141230	381	1975	1984	9
45	*No site*	.	.	.	.	.	.	.
46	WD174_0	West Coast	5863640	2381860	60	1975	1983	8
47	RO2052_4	Central North Island	6317010	2775330	430	1985	1990	5
48	FR202_3	Central North Island	6335015	2819550	231	1993	2000	7
49	FR216_1	Canterbury	5771035	2479825	22	1994	2000	6
50	FR217_1	Central North Island	6297170	2821140	495	1994	2001	7
51	FR217_2	Central North Island	6305480	2755850	240	1994	2001	7
52	FR217_3	Central North Island	6337980	2848110	215	1994	2001	7
53	FR259_2	Central North Island	6258750	2749620	443	1995	2004	9
54	FR259_3	Australia	.	.	.	.	.	.
55	FR260_1	Central North Island	6317930	2835925	332	1995	2002	7
56	FR260_3	Central North Island	6333900	2823050	117	1995	2002	7
57	*No site*	.	.	.	.	.	.	.
58	FR307_2	Central North Island	6312810	2752900	224	1997	2004	7
59	FR307_3	Australia	.	.	.	1997	2005	8
60	FR259_1	Central North Island	6317930	2835925	332	1995	2004	9
61	FR123_1	Central North Island	6335810	2830550	61	1990	1997	7
62	FR123_4	Central North Island	6320510	2761580	379	1990	1997	7
63	FR124_1	Central North Island	6286420	2786900	412	1990	1999	9
64	FR124_4	Otago/Southland	5464490	2292580	27	1990	1999	9
65	FR305_11	Central North Island	6330902	2819285	320	1997	2005	8
66	FR305_12	Central North Island	6330902	2819285	320	1997	2005	8
67	FR305_21	Auckland	6512183	2623611	35	1997	2005	8
68	FR305_22	Auckland	6512183	2623611	35	1997	2005	8
69	FR354_1	Central North Island	6333449	2824785	100	1998	2005	7
70	FR354_2	Central North Island	6321673	2782113	390	1998	2005	7
71	FR353_1	Central North Island	6330562	2819078	320	1999	2006	7
72	FR353_2	Central North Island	6323251	2777453	615	1999	2006	7
73	FR353_3	Auckland	6496483	2634353	30	1999	2006	7
74	FR399_1	Northland	6318025	2935379	270	2000	2007	7
75	FR399_2	Central North Island	6325964	2770215	470	2000	2007	7
76	FR399_3	Central North Island	6332645	2826581	85	2000	2007	7

Figure 17 – Results of clustering using a threshold analysis technique with the threshold relaxed to 0.72, where colours indicate cluster membership.

4.3 Connectivity of the proximity matrix

A proximity matrix is strongly connected if, and only if, all off-diagonal entries are greater than zero in the matrix $B = A + A^2 + \ldots + A^{n-1}$, where A is the adjacency matrix of the proximity matrix (Aldous and Wilson, 2000). The B-matrix is shown in Figure 18. The lack of connectivity between sites is highlighted by a majority of the B-matrix containing zero entries (non-coloured areas).

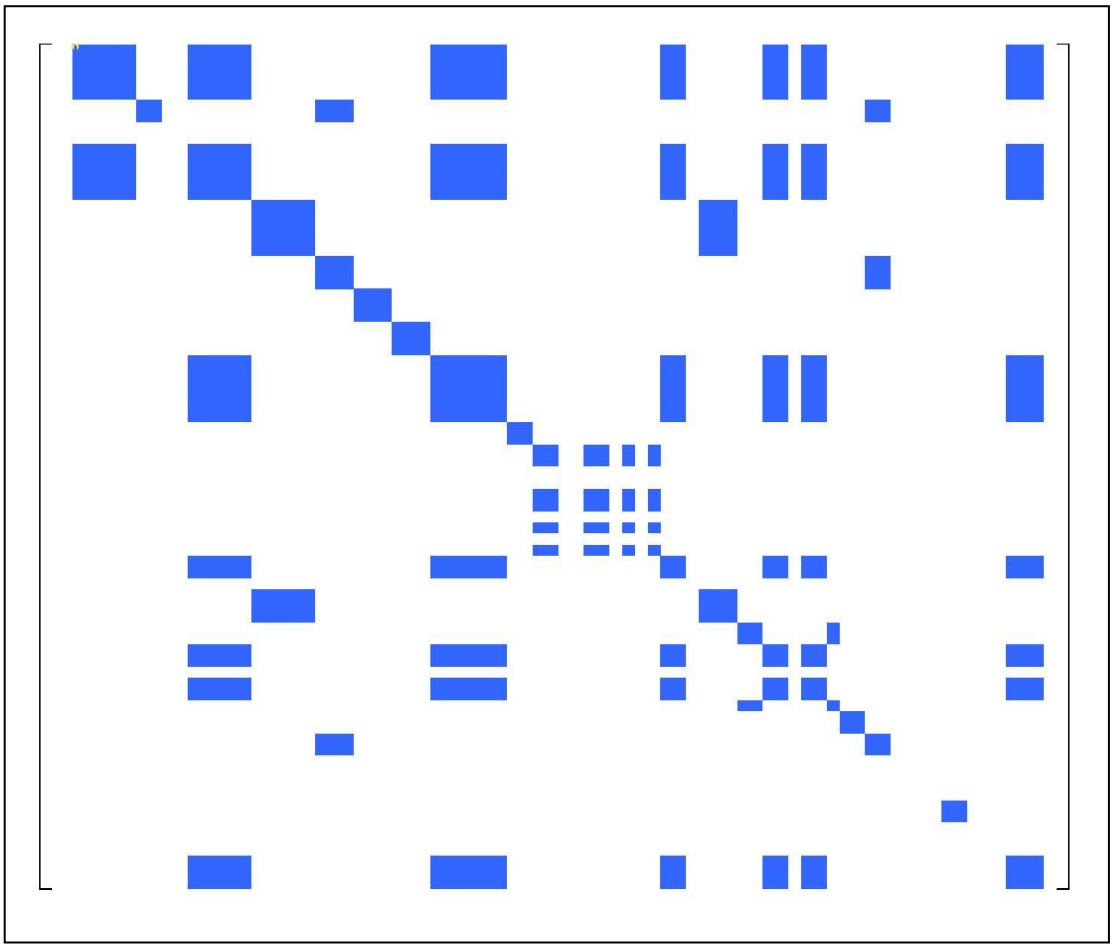


Figure 18 - Schematic of the B-matrix: calculated as B = A + A2 + A3 + … + An-1, where A is the adjacency matrix of the proximity matrix. The blocks of colour represent off-diagonal entries greater than zero. A strongly connected proximity matrix would produce a B-matrix with all off-diagonal entries coloured (i.e. greater than zero).

5. DISCUSSION

Construction of a proximity matrix highlighted an issue of low connectivity between sites, with only 148 out of 2850 cells in the (triangular portion of the 76 x 76) matrix populated. The lack of connectivity exists due to sites having insufficient families in common. Given that the trial sites were established over such a long time period, it is likely that each trial series was planted with families that were considered potential "winners" at the time. As subsequent trial series were established, opinions and objectives changed and the group of "winners" appears to have been adjusted to fit with the thinking of the day.

The low connectivity between sites meant that the correlation matrix contained many missing values, rendering it impossible to use many common analytical techniques for grouping sites. Instead, some basic concepts from Graph Theory were used to develop a threshold analysis technique for clustering the sites. This technique proved to be analytically and computationally simple. In addition, the method was not impacted by the large number of missing values, and shows a lot of promise for other applications where the aim is to group objects based on proximity data.

For future trial establishment it is imperative that sufficient connectivity with previous trials is maintained to allow comparisons at a later date. It is important not to let trends of the day override the ability for comparison, by making any one trial series internally cohesive at the same time as being externally isolated from other trial series. For any trial series intended to be useful for a decision on regionalisation, there must necessarily be a dedicated set of families linking this trial series to all other trial series, even if the linking families do not contribute to the main purpose of the trial series. In this way, external isolation should instead be manifested in the partitioning of sites based on genotype x environment interaction, as opposed to the genotype effect.

In addition, it is important for future trial series to be established over a larger geographic range. The concentration of sites around the central North Island does little to assist the task of segregating sites into breeding regions. A concerted effort

must be made to ensure that other parts of New Zealand, particularly the South Island are included as well as maintaining a sufficient number of trial sites in the central North Island.

A factor which further detracted from the ability to determine appropriate groupings of sites has been the availability of accurate climatic data. It was often the case that weather stations located optimally relative to trial locations were hard to find, meaning an average distance to a weather station of 27 km and some weather stations as far as 103 km from the trial site. When weather stations were selected, it was not uncommon for recording of certain data to cease prior to, or during the life of, the trial. In these cases it was then necessary to supplement the data from one station with data from a second, and in some cases a third, station. Often weather stations simply did not record certain climatic variables.

The collection of quality climatic data is a tough issue to overcome. The ideal situation would be to have weather stations located at each trial. However, it is acknowledged that this is not likely to be an economic option. Establishing and maintaining communication with meteorological agencies may be a means of signalling an interest in the stations and associated data. As an alternative, Landcare Research is moving toward the derivation of climatic variables across a topographic surface using their Digital Elevation Model. This may provide a future method for interpolating more precise climatic data at the actual trial location. Further, soil properties information may also be available in a similar form.

Despite these complications, all but one of the sites was connected in some way to at least one other site. This meant there was sufficient information available to produce a multi-dimensional scaling plot to represent the similarity of sites to each other. Unfortunately, no obvious pattern was noted in the multi-dimensional scaling plot, either in terms of site geography or available environmental variables.

It is acknowledged that this study does not attempt to analyse genotype by environment interactions at the seedlot level, which from a practical point of view would have given a useful result to forestry companies and seed producers. It was decided that many forestry companies will plant a mix of genetic material which

changes from year to year, making it difficult to define seedlots that warrant investigation. However, this mix of genetic material provides an inherent buffer against the effect of the environment on an average performance basis. In addition, those companies that are aggressive in their selection of genetic material will find knowledge presented here of family performance helpful for identifying parents for specific crosses.

Genotype performance across sites was compared using a level plot. Matheson and Raymond (1984) pointed out the dearth of knowledge regarding the across site performance of *P. radiata*. Figure 6 provides evidence of *P. radiata* families that perform consistently in terms of growth across most or all sites on which they are established. Although families did change rank between sites, further analysis of genetic correlations between sites showed that ranking reversals were infrequent. This result agrees with results reported by Matheson and Raymond (1984), who noted some reversals of ranking, but found rank changes across Australian sites to be more common.

In the RPBC data, there were also families that were very interactive. This leaves New Zealand breeders and forest owners in a strong position, as they are able to choose between a number of breeding strategies. The more favoured strategies in the literature appear to be exclusion of interactive genotypes and exploitation of interactive genotypes by grouping environments into breeding regions where groups of genotypes will be most successful. The former option has the disadvantage of removing families that may be the best performers at some sites. The latter option is dependent on the ability to segregate environments based on one or a number of dictating factors.

The main driver of genotype x environment interactions in this study was found to be extreme maximum temperatures. Extreme temperatures have been found to significantly affect foliar conductance, and consequently the acquisition of biomass (Bassow et al., 1994). Bassow et al. (1994) exposed seedlings of three temperate species to temperatures of 45°C for one day, and noted significantly decreased biomass 35 and 105 days later. The highest extreme temperature across all of the RPBC trials was 42.4°C, and there were numerous temperatures recorded greater than

35°C in the summer months. And, given the young age of the trials at measurement, the reduction in foliar conductance and the genotypic response of a family to this reduction, may provide a feasible explanation for the changes in rank across sites.

If extreme maximum temperature is important, it may make regionalisation a difficult objective. It is acknowledged that mean global temperatures are currently increasing (Schneider, 1989, Gates et al., 1992, Mitchell and Gregory, 1992). However extreme weather events are more difficult to predict, but, there are indications that the occurrence of extreme temperatures will become more common (Mearns et al., 1984, Rind et al., 1989). It is suggested here that a regional breeding programme based on extreme maximum temperatures could be nearly impossible to implement.

Another option for capturing the response of families to extreme maximum temperatures would be to use predicted regional temperature extremes as the basis for deployment. In this way, planting stock could be determined based on extreme temperature patterns predicted for each forest area. Difficulties with predicting temperature extremes for entire rotations would also make implementation of this strategy onerous. In addition, there is potential for forest areas to change regions frequently, as latest predictions of climate behaviour become available.

The effect of altitude was also found to be significant (at $\alpha = 0.1$ significance level) as a factor influencing genotype x environment interactions. Altitude has been found to be important for genotype x environment interactions of *P. radiata* in Australia (Raymond and Henson, 2009, Wu and Matheson, 2005). Raymond and Henson (2009) highlighted the strong association of altitude with a number of other bioclimatic variables (including maximum temperatures), and implicated temperature and moisture availability as driving factors across the altitude range. Sands and Mulligan (1990) have shown that a decline in water availability (increased vapour-saturation deficit) will decrease foliar conductance. This will result in a corresponding reduction in transpiration rate, and subsequently, reduced growth (McNaughton and Black, 1973).

The regular occurrence of altitude (and its associated bioclimatic characteristics) in the results of genotype x environment interaction studies, would lead one to believe

that it is an important factor in the performance of families at different sites. Fortunately, altitude is a much more stable site characteristic than many other environmental variables (such as extreme temperatures). Consequently, it provides a much more feasible basis for partitioning forest areas into regions for breeding and/or deployment.

Most forest areas in New Zealand encompass a wide range of altitudes. It might, therefore, be possible to maintain a centralised breeding population and deploy families according to regional altitudinal ranges. However, altitude may be acting as a proxy for the true effect(s) acting upon family performance. In this case, response to altitude may differ in certain parts of New Zealand depending on the action of some bioclimatic variables across the altitude range. If this is so, then a more regional approach to breeding may be warranted.

6. CONCLUSIONS

An objective of this study was to identify environmental variables driving genotype x environment interactions. Extreme maximum temperatures ($\alpha = 0.05$) and altitude ($\alpha = 0.1$) were both found to have significant effects on genetic correlations between sites. The effect of these two climatic variables may be proxies for the physiological influence of foliar conductance and moisture availability.

A second objective was to identify interacting sites and *P.radiata* genotypes. The exact delineation of sites into regions is unclear following this study. However, there is certainly evidence of genotype x environment interactions, with a number of families displaying an unstable performance across sites. Low connectivity between sites has made partitioning sites into regions difficult, but can be improved through closer links between past, present, and future trial series. In addition, there is a need to improve the coverage and recording of climatic information.

The two core methods of analysis used in this study, multiple regression on resemblance matrices and threshold clustering analysis, have both proved to be conceptually and computationally simple. The use of these techniques is encouraged for other quantitative genetics applications. The latter of the two in particular, appears to be new to the field of quantitative genetics, and allowed the issue of missing data to be easily side-stepped.

Despite not finding useful groupings of sites, the discovery of some extremely interactive genotypes does lead this author to conclude that the partitioning of breeding and/or deployment efforts should not be discounted. There is an opportunity to take advantage of the natural variability of *P. radiata* families, by establishing those that are most suited to the environmental conditions at a site. However, the exact partitioning of the breeding programme cannot be described without better cohesion between trial series, and more accurate bioclimatic data.

7. REFERENCES

Ades, P. K. & Garnier-Gere, P.-H. (1997) Making Sense of Provenance x Environment Interaction in *Pinus radiata*. *Proceedings of IUFRO '97 Genetics of Radiata Pine*, pp. 113-119. Rotorua, New Zealand, 1-4 December 1997.

Aldous, J. M. & Wilson, R. J. (2000) *Graphs and Applications: an introductory approach.* Springer, London.

Amer, P. R., Kemp, R. A. & Smith, C. (1992) Genetic Differences among the Predominant Beef Cattle Breeds in Canada - an Analysis of Published Results. *Canadian Journal of Animal Science,* 72, 759-771.

Bassow, S. L., McConnaughay, K. D. M. & Bazzaz, F. A. (1994) The Response of Temperate Tree Seedlings Grown in Elevated CO_2 to Extreme Temperature Events. *Ecological Applications,* 4, 593-603.

Burdon, R. D. (1971) Clonal Repeatabilities and Clone-Site Interactions in *Pinus radiata*. *Silvae Genetica,* 20, 33-39.

Burdon, R. D. (1975) Compression Wood in *Pinus radiata* Clones on Four Different Sites. *New Zealand Journal of Forestry Science,* 5, 152-164.

Burdon, R. D. (1977) Genetic Correlation as a Concept for Studying Genotype-Environment Interaction in Forest Tree Breeding. *Silvae Genetica,* 26, 168-175.

Burdon, R. D., Firth, A., Low, C. B. & Miller, M. A. (1997) Native provenances of *Pinus radiata* in New Zealand: performance and potential. *New Zealand Journal of Forestry,* 41, 32-36.

Carson, M. J., Burdon, R. D., Carson, S. D., Firth, A., Shelbourne, C. J. A. & Vincent, T. G. (1990) Realising Genetic Gains in Production Forests. *Breeding and Genetic Resources. Proceedings of Joint Meetings of Western Forest Genetics Association and IUFRO Working Parties on Douglas-fir, Lodgepole pine, Sitka, and Abies spp.* Olympia, Washington, U.S.A.

Carson, S. D. (1991) Genotype x Environment Interaction and Optimal Number of Progeny Test Sites for Improving *Pinus radiata* in New Zealand. *New Zealand Journal of Forestry Science,* 21, 32-49.

Douglas, M. E. & Endler, J. A. (1982) Quantitative Matrix Comparisons in Ecological and Evolutionary Investigations. *Journal of Theoretical Biology,* 99, 777-795.

Dow, M. M. & Cheverud, J. M. (1985) Comparison of distance matrices in studies of population structure and genetic microdifferentiation: Quadratic assignment. *American Journal of Physical Anthropology,* 68, 367-373.

ESRI (2006) ArcMAP 9.2. Environmental Systems Research Institute, Redlands, California.

Falconer, D. S. (1952) The Problem of Environment and Selection. *The American Naturalist,* 86, 293-298.

Gates, W. L., Mitchell, J. F. B., Boer, G. J., Cubasch, U. & Meleshko, V. P. (1992) Climate modelling, climate prediction and model validation. *Climate Change 1992: the supplementary report to the IPCC scientific assessment* (eds T. J. Houghton, B. A. Callander & S. K. Varney), pp. 97-134. Cambridge University Press, Cambridge, England.

Gordon, A. D. (1999) *Classification.* Chapman & Hall/CRC, Boca Raton.

Hodge, G. R. & White, T. L. (1992) Genetic Parameter Estimates for Growth Traits at Different Ages in Slash Pine and Some Implications for Breeding. *Silvae Genetica,* 41, 252-262.

Hubert, L. J. & Golledge, R. G. (1981) A Heuristic Method for the Comparison of Related Structures. *Journal of Mathematical Psychology,* 23, 214-226.

Jayawickrama, K. J. S., Carson, M. J., Jefferson, P. A. & Firth, A. (1997) Development of the New Zealand Radiata Pine Breeding Population. *Proceedings of IUFRO '97 Genetics of Radiata Pine*, pp. 217-225. Rotorua, New Zealand, 1-4 December 1997.

Jinks, J. L. & Mather, K. (1955) Stability in Development of Heterozygotes and Homozygotes. *Proceedings of the Royal Society of London Series B-Biological Sciences,* 143, 561-578.

Johnson, G. R. & Burdon, R. D. (1990) Family-Site Interaction in *Pinus radiata*: Implications for Progeny Testing Strategy and Regionalised Breeding in New Zealand. *Silvae Genetica,* 39, 55-62.

Kumar, S. (2004) Genetic parameter estimates for wood stiffness, strength, internal checking, and resin bleeding for radiata pine. *Canadian Journal of Forest Research,* 34, 2601-2610.

Kumar, S., Burdon, R., Stovold, G. & Gea, L. (2008) Implications of selection history on genetic architecture of growth, form, and wood-quality traits in Pinus radiata. *Canadian Journal of Forest Research,* 38, 2372.

Legendre, P., Lapointe, F. J. & Casgrain, P. (1994) Modeling Brain Evolution from Behavior - a Permutational Regression Approach. *Evolution,* 48, 1487-1499.

Lichstein, J. W. (2007) Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology,* 188, 117-131.

Mantel, N. (1967) The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Res,* 27, 209-220.

Matheson, A. C. & Raymond, C. A. (1984) The Impact of Genotype X Environment Interactions on Australian *Pinus radiata* Breeding Programs. *Australian Forest Research,* 14, 11-25.

Mathur, P. K. & Horst, P. (1994) Methods for Evaluating Genotype-Environment Interactions Illustrated by Laying Hens. *Journal of Animal Breeding and Genetics-Zeitschrift Fur Tierzuchtung Und Zuchtungsbiologie,* 111, 265-288.

McKimmy, M. D. & Campbell, R. K. (1982) Genetic variation in the Wood Density and Ring Width Trend in Coastal Douglas-fir. *Silvae Genetica,* 31, 43-51.

McNaughton, K. G. & Black, T. A. (1973) A Study of Evapotranspiration from a Douglas Fir Forest Using the Energy Balance Approach. *Water Resources Research,* 9, 1579-1590.

Mearns, L. O., Katz, R. W. & Schneider, S. H. (1984) Extreme High-Temperature Events: Changes in their probabilities with Changes in Mean Temperature. *Journal of Climate and Applied Meteorology,* 23, 1601-1613.

Milligan, G. W. & Cooper, M. C. (1985) An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika,* 50, 159-179.

Mitchell, J. F. B. & Gregory, J. M. (1992) Climatic consequences of emissions and a comparison of IS92a and SA90. *Climate Change 1992: the supplementary report to the IPCC scientific assessment* (eds T. J. Houghton, B. A. Callander & S. K. Varney), pp. 171-175. Cambridge University Press, Cambridge, England.

Mulder, H. A. & Bijma, P. (2005) Effects of genotype x environment interaction on genetic gain in breeding programs. *Journal of Animal Science,* 83, 49-61.

Owino, F. (1977a) Genotype-X Environment Interaction and Genotypic Stability in loblolly pine .IV. Correlation Studies. *Silvae Genetica,* 26, 176-179.

Owino, F. (1977b) Genotype x Environment Interaction and Genotypic Stability in loblolly pine. II. Genotypic stability comparisons. *Silvae Genetica,* 26, 21-26.

Owino, F., Kellison, R. C. & Zobel, B. J. (1977) Genotype-X Environment Interaction and Genotypic Stability in loblolly pine .V. Effects of Genotype-X Environment Interaction on Genetic Variance Component Estimates and Gain Predictions. *Silvae Genetica,* 26, 131-134.

Owino, F. & Zobel, B. (1977a) Genotype X Environment Interaction and Genotypic Stability in loblolly pine .III. Heterosis and Heterosis X Environment Interaction. *Silvae Genetica,* 26, 114-116.

Owino, F. & Zobel, B. (1977b) Genotype x Environment Interaction and Genotypic Stability in loblolly pine. I. General introduction and description of the experiment. *Silvae Genetica,* 26, 18-21.

Raymond, C. A. & Henson, M. (2009) Genotype by Environment Interaction in *Pinus radiata* within NSW. *2nd Australasian Forest Genetics Conference, 20th-22nd April 2009*. Fremantle, Australia.

Rind, D., Goldberg, R. & Ruedy, R. (1989) Change in climate variability in the 21st century. *Climatic Change,* 14, 5-37.

Sands, R. & Mulligan, D. R. (1990) Water and nutrient dynamics and tree growth. *Forest Ecology and Management,* 30, 91-111.

Schneider, S. H. (1989) The Greenhouse Effect: Science and Policy. *Science,* 243, 771.

Schnell, G. D., Watt, D. J. & Douglas, M. E. (1985) Statistical Comparison of Proximity Matrices - Applications in Animal Behavior. *Animal Behaviour,* 33, 239-253.

Searle, S. R. (1982) *Matrix algebra useful for statistics.* John Wiley & Sons, Inc., New York.

Shelbourne, C. J. A., Burdon, R. D., Carson, S. D., Firth, A. & Vincent, T. G. (1986) *Development plan for radiata pine breeding.* Forest Research Institute, Rotorua, New Zealand.

Shelbourne, C. J. A., Carson, M. J. & Wilcox, M. D. (1989) New techniques in the genetic improvement of radiata pine. *Commonwealth Forestry Review,* 68, 191-201.

Shelbourne, C. J. A. & Low, C. B. (1980) Multi-trait Index Selection and Associated Genetic Gains of *Pinus radiata* Progenies at Five Sites. *New Zealand Journal of Forestry Science,* 10, 307-324.

Simons, A. M. & Roff, D. A. (1996) The effect of a variable environment on the genetic correlation structure in a field cricket. *Evolution,* 50, 267-275.

Smouse, P. E., Long, J. C. & Sokal, R. R. (1986) Multiple-Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence. *Systematic Zoology,* 35, 627-632.

Sutton, W. R. J. (2007) The quality of the present harvest is the result of early management decisions. *New Zealand Journal of Forestry,* 51, 38-39.

Thulin, I. J. (1957) Application of Tree Breeding to Forestry in New Zealand. *New Zealand Journal of Forestry,* 7, 41-46.

Walker, J. C. F. (2007) Clausewitz on Forestry. *New Zealand Journal of Forestry,* 51, 34-37.

Wu, H. X. & Matheson, C. A. (2005) Genotype by Environment Interactions in an Australia-Wide Radiata Pine Diallel Mating Experiment: Implications for Regionalized Breeding. *Forest Science,* 51, 29.

Yamada, Y. (1962) Genotype by Environment Interaction and Genetic Correlation of the Same Trait under Different Environments. *Japanese Journal of Genetics,* 37, 498-509.

Zobel, B. & Talbert, J. (1984) *Applied Forest Tree Improvement.* John Wiley and Sons, New York.

8. APPENDIX A – LIST OF TRIAL LOCATIONS

Site number	Trial id	Wood region	Northing	Easting	Altitude	Year of Establishment	Year of Measurement	Age at Measurement
1	RO944_8	Central North Island	6318820	2832690	415	1968	1978	10
2	WN212_0	Hawkes Bay	6158726	2796271	780	1969	1979	10
3	RO947_2	Central North Island	6262553	2781916	466	1969	1979	10
4	AK290_0	Northland	6504000	2629358	47	1971	1984	13
5	Sd228_0	Otago/Southland	5453769	2286340	227	1971	1984	13
6	RO1015_1	Central North Island	6258790	2781600	499	1972	1981	9
7	RO320_16	Central North Island	6298060	2816510	554	1972	1982	10
8	RO1015_2	Central North Island	6258116	2781760	518	1972	1982	10
9	RO320_15	Central North Island	6298370	2815720	548	1972	1982	10
10	AK622_1	Northland	6523042	2618011	27	1975	1983	8
11	AK623_1	Auckland	6431268	2710664	80	1975	1983	8
12	NN330_1	Nelson/Marlborough	5947614	2495849	590	1975	1983	8
13	RO664_1	Central North Island	6281630	2801340	563	1975	1984	9
14	SD415_0	Otago/Southland	5467683	2271120	520	1975	1983	8
15	RO1836_0	Central North Island	6352370	2821040	280	1981	1989	8
16	RO1884_3	Central North Island	6297860	2816460	554	1983	1990	7
17	RO1838_0	Central North Island	6248133	2781409	580	1981	1985	4
18	RO1884_2	Central North Island	6307253	2773875	259	1983	1989	6
19	FR171_3	Central North Island	6334665	2827645	94	1992	2000	8
20	AK1061_1	Northland	6536939	2611689	100	1987	1995	8
21	AK1061_2	Northland	6603206	2587078	100	1987	1993	6
22	RO2111_1	Central North Island	6281100	2801220	565	1987	1995	8
23	FR38_1	Northland	6537300	2609610	81	1988	1996	8
24	FR38_2	Central North Island	6321505	2761135	372	1988	1996	8
25	FR38_3	Central North Island	6362670	2808150	98	1988	1995	7
26	FR69_1	Northland	6537300	2609610	81	1989	1998	9
27	FR69_2	Central North Island	6322545	2759615	357	1989	1997	8
28	FR69_3	Central North Island	6312985	2834650	296	1989	1996	7
29	FR170_1	Northland	6492870	2635580	20	1992	2001	9
30	FR170_2	Central North Island	6327278	2772802	509	1992	2000	8
31	FR170_3	Central North Island	6312709	2835293	300	1992	2000	8
32	FR203_1	Hawkes Bay	6212530	2829150	451	1993	2001	8
33	FR203_2	Australia	.	.		1993	2000	7
34	FR203_3	Central North Island	6314034	2830097	420	1993	2001	8
35	RO1804_0	Central North Island	6344950	2837800	20	1980	1987	7
36	RO664_13	Central North Island	6281680	2801266	561	1980	1987	7
37	AK622_2	.	.	.	.	1975	1983	8
38	AK623_2	Auckland	6431390	2710715	100	1975	1983	8
39	*No site*	.	.	.	.	.	.	.
40	*No site*	.	.	.	.	.	.	.

Site number	Trial id	Wood region	Northing	Easting	Altitude	Year of Establishment	Year of Measurement	Age at Measurement
41	RO320_25	Central North Island	6298060	2816510	554	1975	1984	9
42	RO663_0	East Coast	6370745	2973785	370	1975	1983	8
43	RO664_2	Central North Island	6281680	2801260	561	1975	1984	9
44	SD413_0	Otago/Southland	5461200	2141230	381	1975	1984	9
45	*No site*	.	.	.	.	.	.	.
46	WD174_0	West Coast	5863640	2381860	60	1975	1983	8
47	RO2052_4	Central North Island	6317010	2775330	430	1985	1990	5
48	FR202_3	Central North Island	6335015	2819550	231	1993	2000	7
49	FR216_1	Canterbury	5771035	2479825	22	1994	2000	6
50	FR217_1	Central North Island	6297170	2821140	495	1994	2001	7
51	FR217_2	Central North Island	6305480	2755850	240	1994	2001	7
52	FR217_3	Central North Island	6337980	2848110	215	1994	2001	7
53	FR259_2	Central North Island	6258750	2749620	443	1995	2004	9
54	FR259_3	Australia	.	.	.	.	.	.
55	FR260_1	Central North Island	6317930	2835925	332	1995	2002	7
56	FR260_3	Central North Island	6333900	2823050	117	1995	2002	7
57	*No site*	.	.	.	.	.	.	.
58	FR307_2	Central North Island	6312810	2752900	224	1997	2004	7
59	FR307_3	Australia	.	.	.	1997	2005	8
60	FR259_1	Central North Island	6317930	2835925	332	1995	2004	9
61	FR123_1	Central North Island	6335810	2830550	61	1990	1997	7
62	FR123_4	Central North Island	6320510	2761580	379	1990	1997	7
63	FR124_1	Central North Island	6286420	2786900	412	1990	1999	9
64	FR124_4	Otago/Southland	5464490	2292580	27	1990	1999	9
65	FR305_11	Central North Island	6330902	2819285	320	1997	2005	8
66	FR305_12	Central North Island	6330902	2819285	320	1997	2005	8
67	FR305_21	Auckland	6512183	2623611	35	1997	2005	8
68	FR305_22	Auckland	6512183	2623611	35	1997	2005	8
69	FR354_1	Central North Island	6333449	2824785	100	1998	2005	7
70	FR354_2	Central North Island	6321673	2782113	390	1998	2005	7
71	FR353_1	Central North Island	6330562	2819078	320	1999	2006	7
72	FR353_2	Central North Island	6323251	2777453	615	1999	2006	7
73	FR353_3	Auckland	6496483	2634353	30	1999	2006	7
74	FR399_1	Northland	6318025	2935379	270	2000	2007	7
75	FR399_2	Central North Island	6325964	2770215	470	2000	2007	7
76	FR399_3	Central North Island	6332645	2826581	85	2000	2007	7

9. APPENDIX B – LIST OF GENETIC CORRELATIONS

Pair	Site 1	Trial id 1	Site 2	Trial id 2	Families in common	Genetic correlation	Standard error	Convergence status
1	1	RO944_8	2	WN212_0	372	0.68	0.09	Positive
2	1	RO944_8	3	RO947_2	373	0.75	0.06	Positive
3	1	RO944_8	4	AK290_0	136	0.56	0.12	Positive
4	1	RO944_8	5	Sd228_0	146	0.60	0.14	Positive
5	1	RO944_8	47	RO2052_4	66	0.50	0.25	Positive
6	2	WN212_0	3	RO947_2	565	0.68	0.08	Positive
7	2	WN212_0	4	AK290_0	203	0.57	0.13	Positive
8	2	WN212_0	5	Sd228_0	217	0.54	0.16	Positive
9	2	WN212_0	29	FR170_1	26	0.84	0.60	Positive
10	2	WN212_0	30	FR170_2	29	0.11	0.32	Positive
11	2	WN212_0	31	FR170_3	29	0.22	0.37	Positive
12	2	WN212_0	32	FR203_1	33	0.20	0.32	Positive
13	2	WN212_0	33	FR203_2	29	0.24	0.34	Positive
14	2	WN212_0	34	FR203_3	20	0.12	0.47	Positive
15	2	WN212_0	47	RO2052_4	103	0.42	0.26	Positive
16	2	WN212_0	48	FR202_3	24	-0.34	0.46	Positive
17	2	WN212_0	55	FR260_1	23	0.45	0.44	Positive
18	2	WN212_0	56	FR260_3	21	0.30	.	Positive
19	3	RO947_2	4	AK290_0	208	0.34	0.12	Positive
20	3	RO947_2	5	Sd228_0	224	0.52	0.13	Positive
21	3	RO947_2	29	FR170_1	28	1.00	.	Boundary
22	3	RO947_2	30	FR170_2	32	0.34	0.23	Positive
23	3	RO947_2	31	FR170_3	32	0.70	0.24	Positive
24	3	RO947_2	32	FR203_1	37	0.59	0.23	Positive
25	3	RO947_2	33	FR203_2	33	0.25	0.26	Positive
26	3	RO947_2	34	FR203_3	23	0.21	0.38	Positive
27	3	RO947_2	47	RO2052_4	109	0.59	0.19	Positive
28	3	RO947_2	48	FR202_3	26	0.31	0.37	Positive
29	3	RO947_2	55	FR260_1	23	0.56	0.36	Positive
30	3	RO947_2	56	FR260_3	21	0.35	0.85	Positive
31	4	AK290_0	5	Sd228_0	268	0.55	0.10	Positive
32	4	AK290_0	47	RO2052_4	49	0.06	0.30	Positive
33	5	Sd228_0	32	FR203_1	20	0.33	0.46	Positive
34	5	Sd228_0	47	RO2052_4	53	0.31	0.36	Positive
35	6	RO1015_1	7	RO320_16	105	0.75	0.11	Positive
36	6	RO1015_1	63	FR124_1	43	0.43	0.34	Positive
37	6	RO1015_1	64	FR124_4	34	0.54	0.35	Positive
38	7	RO320_16	63	FR124_1	43	0.84	0.33	Positive
39	7	RO320_16	64	FR124_4	34	0.79	0.34	Positive
40	8	RO1015_2	9	RO320_15	27	1.00	.	Boundary
41	8	RO1015_2	37	AK622_2	25	1.00	.	Boundary
42	8	RO1015_2	38	AK623_2	25	1.00	.	Boundary
43	8	RO1015_2	41	RO320_25	25	1.00	.	Boundary
44	8	RO1015_2	42	RO663_0	25	1.00	.	Boundary
45	8	RO1015_2	43	RO664_2	25	1.00	.	Boundary

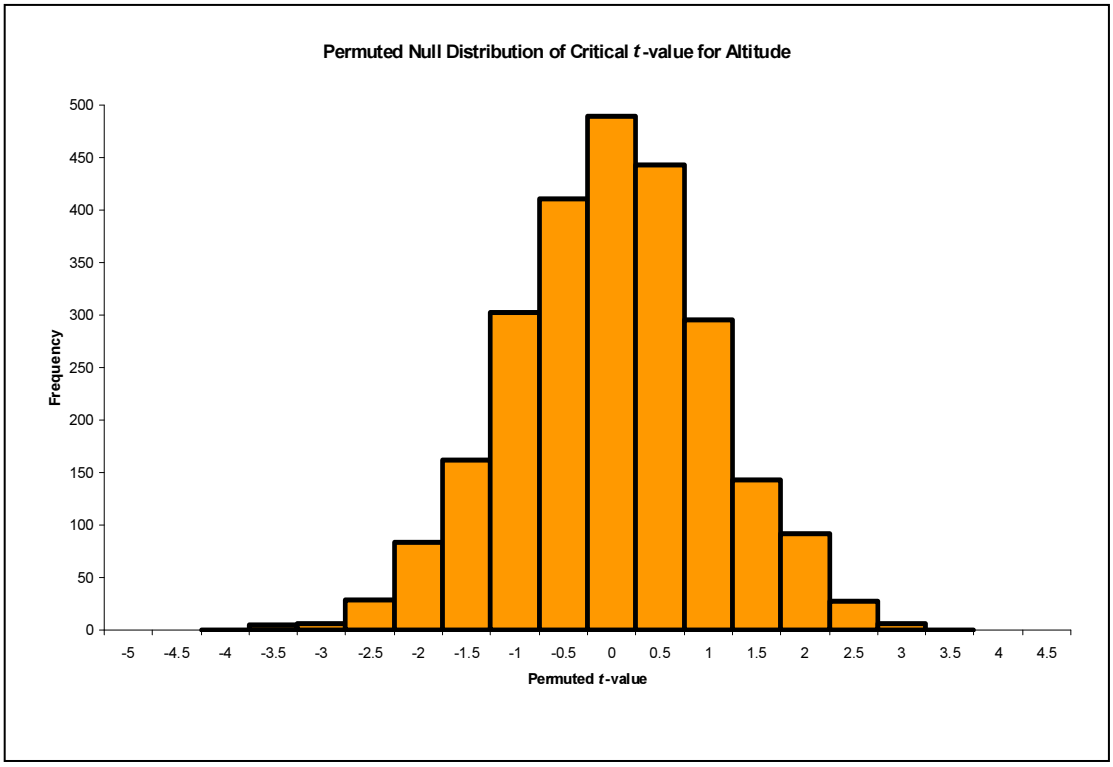
Pair	Site 1	Trial id 1	Site 2	Trial id 2	Families in common	Genetic correlation	Standard error	Convergence status
46	8	RO1015_2	44	SD413_0	25	-1.00	.	Boundary
47	8	RO1015_2	46	WD174_0	25	1.00	.	Boundary
48	9	RO320_15	37	AK622_2	25	1.00	.	Boundary
49	9	RO320_15	38	AK623_2	25	1.00	.	Boundary
50	9	RO320_15	41	RO320_25	25	1.00	.	Boundary
51	9	RO320_15	42	RO663_0	25	1.00	.	Boundary
52	9	RO320_15	43	RO664_2	25	1.00	.	Boundary
53	9	RO320_15	44	SD413_0	25	1.00	.	Boundary
54	9	RO320_15	46	WD174_0	25	1.00	.	Boundary
55	10	AK622_1	11	AK623_1	100	0.36	0.19	Positive
56	10	AK622_1	12	NN330_1	99	0.44	0.14	Positive
57	10	AK622_1	13	RO664_1	99	0.40	0.12	Positive
58	10	AK622_1	14	SD415_0	101	0.54	0.13	Positive
59	10	AK622_1	32	FR203_1	26	0.94	0.24	Positive
60	10	AK622_1	33	FR203_2	23	0.74	0.29	Positive
61	11	AK623_1	12	NN330_1	106	0.47	0.19	Positive
62	11	AK623_1	13	RO664_1	107	0.56	0.15	Positive
63	11	AK623_1	14	SD415_0	100	0.59	0.18	Positive
64	11	AK623_1	30	FR170_2	20	-0.21	0.45	Positive
65	11	AK623_1	31	FR170_3	20	-0.14	0.50	Positive
66	11	AK623_1	32	FR203_1	27	0.33	0.46	Positive
67	11	AK623_1	33	FR203_2	24	1.00	.	Boundary
68	12	NN330_1	13	RO664_1	106	0.70	0.09	Positive
69	12	NN330_1	14	SD415_0	99	0.59	0.13	Positive
70	12	NN330_1	30	FR170_2	20	-0.04	0.35	Positive
71	12	NN330_1	31	FR170_3	20	0.11	0.38	Positive
72	12	NN330_1	32	FR203_1	27	0.02	0.36	Positive
73	12	NN330_1	33	FR203_2	24	0.34	0.34	Positive
74	13	RO664_1	14	SD415_0	99	0.39	0.13	Positive
75	13	RO664_1	30	FR170_2	20	-0.12	0.29	Positive
76	13	RO664_1	31	FR170_3	20	-0.06	0.31	Positive
77	13	RO664_1	32	FR203_1	27	0.53	0.27	Positive
78	13	RO664_1	33	FR203_2	24	0.27	0.31	Positive
79	14	SD415_0	32	FR203_1	26	0.20	0.38	Positive
80	14	SD415_0	33	FR203_2	23	0.39	0.37	Positive
81	15	RO1836_0	16	RO1884_3	170	0.76	0.09	Positive
82	15	RO1836_0	17	RO1838_0	171	0.74	0.09	Positive
83	15	RO1836_0	18	RO1884_2	170	0.68	0.10	Positive
84	15	RO1836_0	19	FR171_3	129	0.63	0.12	Positive
85	15	RO1836_0	50	FR217_1	20	0.34	0.50	Positive
86	15	RO1836_0	51	FR217_2	20	0.48	0.37	Positive
87	16	RO1884_3	17	RO1838_0	170	0.57	0.11	Positive
88	16	RO1884_3	18	RO1884_2	170	0.81	0.09	Positive
89	16	RO1884_3	19	FR171_3	129	0.57	0.13	Positive
90	16	RO1884_3	50	FR217_1	20	0.59	0.49	Positive

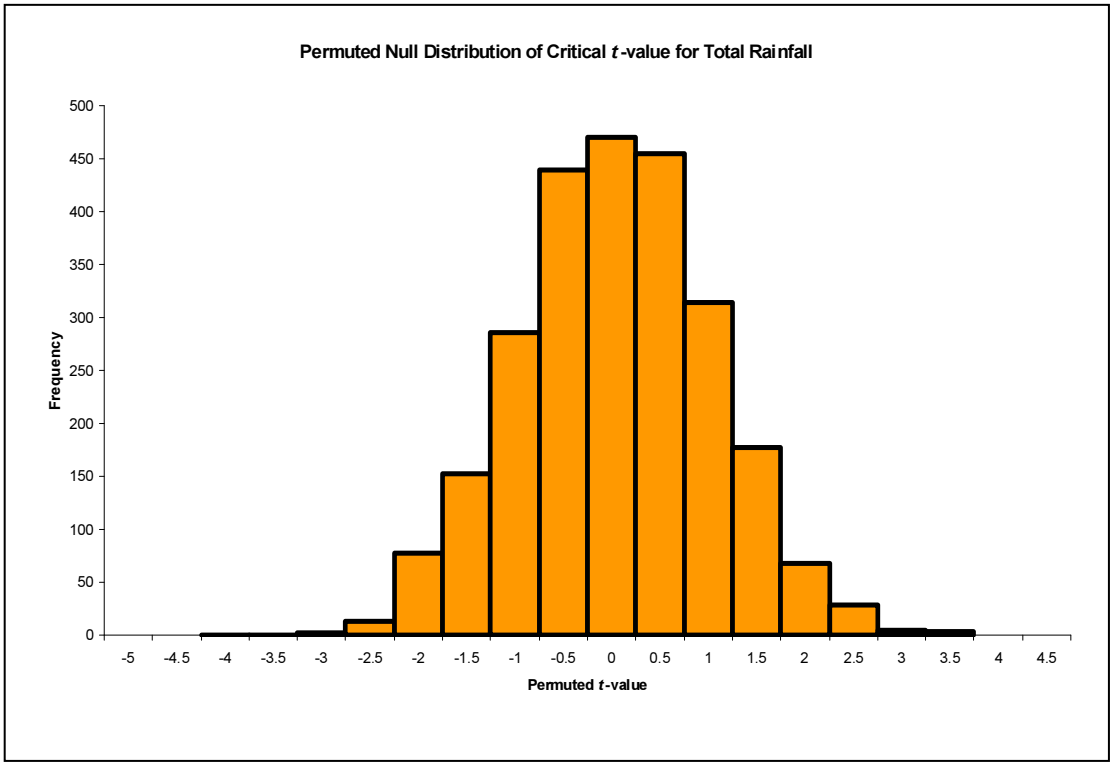
Pair	Site 1	Trial id 1	Site 2	Trial id 2	Families in common	Genetic correlation	Standard error	Convergence status
91	16	RO1884_3	51	FR217_2	20	0.38	0.39	Positive
92	17	RO1838_0	18	RO1884_2	170	0.54	0.12	Positive
93	17	RO1838_0	19	FR171_3	129	0.45	0.14	Positive
94	17	RO1838_0	50	FR217_1	20	1.00	.	Boundary
95	17	RO1838_0	51	FR217_2	20	0.20	0.46	Positive
96	18	RO1884_2	19	FR171_3	129	0.39	0.15	Positive
97	18	RO1884_2	50	FR217_1	20	0.37	0.61	Positive
98	18	RO1884_2	51	FR217_2	20	0.25	0.44	Positive
99	20	AK1061_1	21	AK1061_2	468	0.73	0.09	Positive
100	20	AK1061_1	22	RO2111_1	468	0.50	0.07	Positive
101	21	AK1061_2	22	RO2111_1	468	0.49	0.10	Positive
102	22	RO2111_1	63	FR124_1	52	-0.75	0.80	Positive
103	22	RO2111_1	64	FR124_4	44	0.37	0.84	Positive
104	23	FR38_1	24	FR38_2	225	0.58	0.10	Positive
105	23	FR38_1	25	FR38_3	225	0.61	0.12	Positive
106	24	FR38_2	25	FR38_3	225	0.76	0.13	Positive
107	26	FR69_1	27	FR69_2	330	0.68	0.06	Positive
108	26	FR69_1	28	FR69_3	330	0.64	0.07	Positive
109	27	FR69_2	28	FR69_3	330	0.81	0.06	Positive
110	29	FR170_1	30	FR170_2	131	0.43	0.26	Positive
111	29	FR170_1	31	FR170_3	131	0.28	0.30	Positive
112	29	FR170_1	32	FR203_1	128	0.60	0.34	Positive
113	29	FR170_1	33	FR203_2	121	0.26	0.34	Positive
114	29	FR170_1	34	FR203_3	54	0	.	Boundary
115	29	FR170_1	48	FR202_3	61	0.63	0.49	Positive
116	29	FR170_1	74	FR399_1	20	0	.	Boundary
117	29	FR170_1	76	FR399_3	20	0.35	0.51	Positive
118	30	FR170_2	31	FR170_3	158	0.74	0.13	Positive
119	30	FR170_2	32	FR203_1	155	0.51	0.15	Positive
120	30	FR170_2	33	FR203_2	146	0.45	0.16	Positive
121	30	FR170_2	34	FR203_3	69	1.00	.	Boundary
122	30	FR170_2	47	RO2052_4	20	1.00	.	Boundary
123	30	FR170_2	48	FR202_3	79	0.69	0.27	Positive
124	30	FR170_2	58	FR307_2	21	0.75	0.45	Positive
125	30	FR170_2	59	FR307_3	21	1.00	.	Boundary
126	30	FR170_2	65	FR305_11	22	0.00	0.00	.
127	30	FR170_2	67	FR305_21	22	0.00	0.00	.
128	30	FR170_2	74	FR399_1	24	0.29	0.29	Positive
129	30	FR170_2	75	FR399_2	23	0.54	0.25	Positive
130	30	FR170_2	76	FR399_3	24	0.49	0.25	Positive
131	31	FR170_3	32	FR203_1	155	0.75	0.13	Positive
132	31	FR170_3	33	FR203_2	146	0.43	0.17	Positive
133	31	FR170_3	34	FR203_3	69	0.91	0.26	Positive
134	31	FR170_3	47	RO2052_4	20	0.96	0.30	Positive
135	31	FR170_3	48	FR202_3	79	1.00	.	Boundary

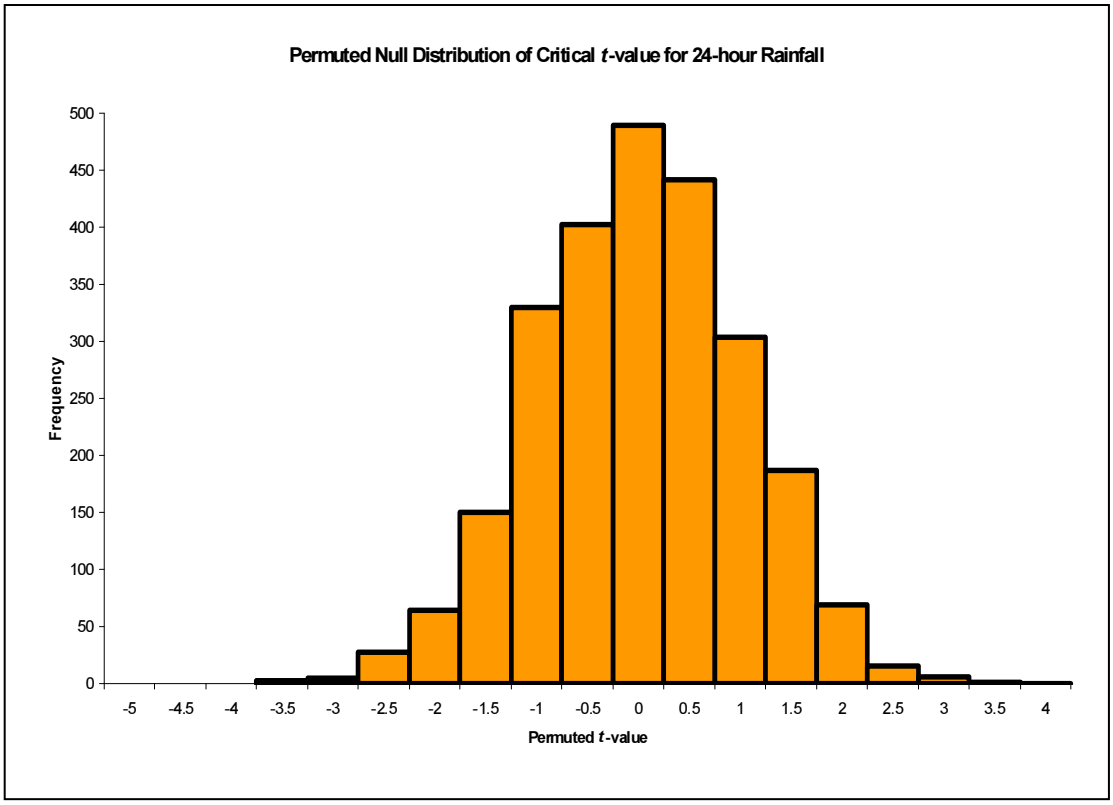
Pair	Site 1	Trial id 1	Site 2	Trial id 2	Families in common	Genetic correlation	Standard error	Convergence status
136	31	FR170_3	58	FR307_2	21	1.00	.	Boundary
137	31	FR170_3	59	FR307_3	21	1.00	.	Boundary
138	31	FR170_3	65	FR305_11	22	0.00	0.00	.
139	31	FR170_3	67	FR305_21	22	0.00	0.00	.
140	31	FR170_3	74	FR399_1	24	0.23	0.29	Positive
141	31	FR170_3	75	FR399_2	23	0.69	0.23	Positive
142	31	FR170_3	76	FR399_3	24	0.89	0.20	Positive
143	32	FR203_1	33	FR203_2	170	0.71	0.15	Positive
144	32	FR203_1	34	FR203_3	103	0.48	0.26	Positive
145	32	FR203_1	47	RO2052_4	20	1.00	.	Boundary
146	32	FR203_1	48	FR202_3	99	0.63	0.26	Positive
147	32	FR203_1	58	FR307_2	21	0.91	0.46	Positive
148	32	FR203_1	59	FR307_3	21	0.65	0.31	Positive
149	32	FR203_1	65	FR305_11	29	0.00	0.00	.
150	32	FR203_1	67	FR305_21	29	0.00	0.00	.
151	32	FR203_1	71	FR353_1	21	0.00	0.00	.
152	32	FR203_1	72	FR353_2	21	0.00	0.00	.
153	32	FR203_1	73	FR353_3	20	0.00	0.00	.
154	32	FR203_1	74	FR399_1	28	0.26	0.29	Positive
155	32	FR203_1	75	FR399_2	28	0.79	0.22	Positive
156	32	FR203_1	76	FR399_3	28	0.87	0.21	Positive
157	33	FR203_2	34	FR203_3	79	0.54	0.29	Positive
158	33	FR203_2	48	FR202_3	86	0.22	0.31	Positive
159	33	FR203_2	58	FR307_2	21	1.00	.	Boundary
160	33	FR203_2	59	FR307_3	21	0.65	0.35	Positive
161	33	FR203_2	65	FR305_11	27	0.00	0.00	.
162	33	FR203_2	67	FR305_21	27	0.00	0.00	.
163	33	FR203_2	74	FR399_1	21	0.00	.	Boundary
164	33	FR203_2	75	FR399_2	21	0.57	0.25	Positive
165	33	FR203_2	76	FR399_3	21	0.70	0.25	Positive
166	34	FR203_3	48	FR202_3	57	0.80	0.41	Positive
167	35	RO1804_0	36	RO664_13	105	0.86	0.12	Positive
168	37	AK622_2	38	AK623_2	25	1.00	.	Boundary
169	37	AK622_2	41	RO320_25	25	1.00	.	Boundary
170	37	AK622_2	42	RO663_0	25	0.44	3.64	Positive
171	37	AK622_2	43	RO664_2	25	-1.00	.	Boundary
172	37	AK622_2	44	SD413_0	25	-0.38	2.16	Positive
173	37	AK622_2	46	WD174_0	25	1.00	.	Boundary
174	38	AK623_2	41	RO320_25	25	0.29	0.63	Positive
175	38	AK623_2	42	RO663_0	25	-0.58	0.76	Positive
176	38	AK623_2	43	RO664_2	25	-0.11	0.84	Positive
177	38	AK623_2	44	SD413_0	25	0.35	0.61	Positive
178	38	AK623_2	46	WD174_0	25	1.00	.	Boundary
179	41	RO320_25	42	RO663_0	25	1.00	.	Boundary
180	41	RO320_25	43	RO664_2	25	1.00	.	Boundary

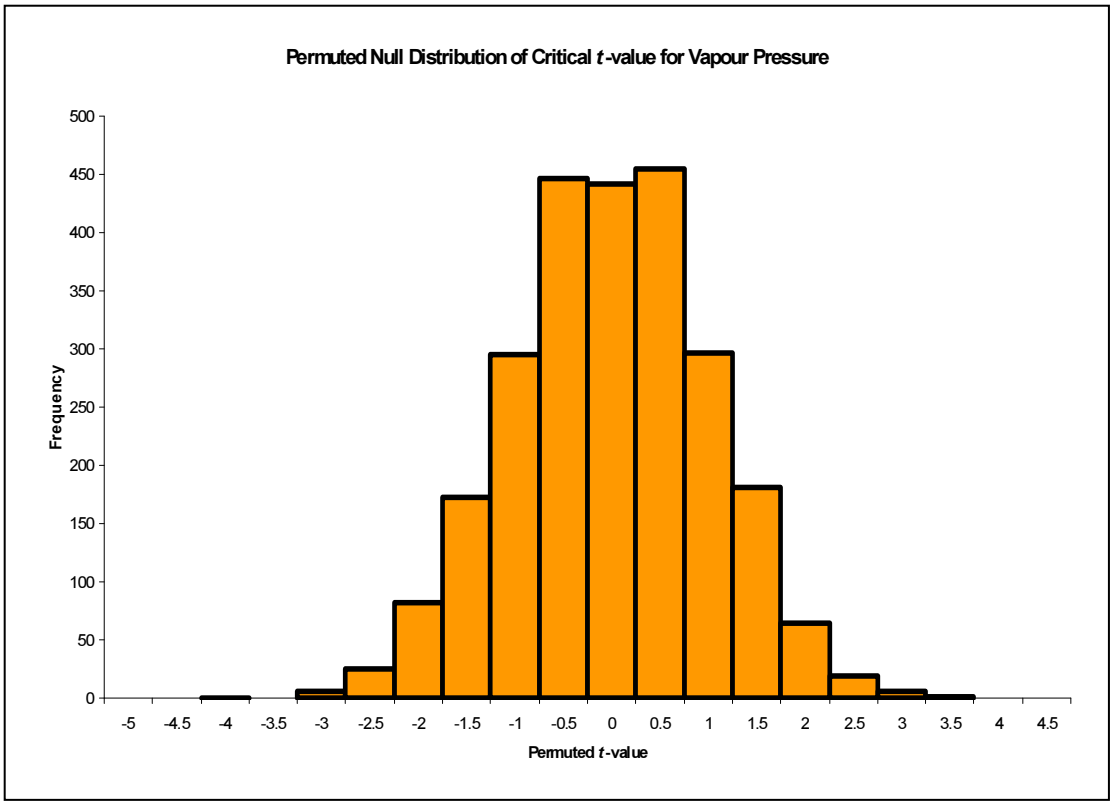
Pair	Site 1	Trial id 1	Site 2	Trial id 2	Families in common	Genetic correlation	Standard error	Convergence status
181	41	RO320_25	44	SD413_0	25	0.87	0.55	Positive
182	41	RO320_25	46	WD174_0	25	1.00	.	Boundary
183	42	RO663_0	43	RO664_2	25	1.00	.	Boundary
184	42	RO663_0	44	SD413_0	25	0.60	0.67	Positive
185	42	RO663_0	46	WD174_0	25	1.00	.	Boundary
186	43	RO664_2	44	SD413_0	25	1.00	.	Boundary
187	43	RO664_2	46	WD174_0	25	1.00	.	Boundary
188	44	SD413_0	46	WD174_0	25	0.68	0.68	Positive
189	48	FR202_3	65	FR305_11	25	0.00	0.00	.
190	48	FR202_3	67	FR305_21	25	0.00	0.00	.
191	50	FR217_1	51	FR217_2	51	0.65	0.31	Positive
192	50	FR217_1	52	FR217_3	47	0.79	0.24	Positive
193	51	FR217_2	52	FR217_3	47	0.89	0.16	Positive
194	53	FR259_2	54	FR259_3	129	0.57	0.12	Positive
195	53	FR259_2	60	FR259_1	129	0.92	0.08	Positive
196	54	FR259_3	60	FR259_1	129	0.43	0.13	Positive
197	55	FR260_1	56	FR260_3	43	0.52	0.79	Positive
198	55	FR260_1	66	FR305_12	35	0.00	0.00	.
199	55	FR260_1	68	FR305_22	35	0.00	0.00	.
200	56	FR260_3	66	FR305_12	35	0.00	0.00	.
201	56	FR260_3	68	FR305_22	35	0.00	0.00	.
202	58	FR307_2	59	FR307_3	58	0.97	0.24	Positive
203	61	FR123_1	62	FR123_4	21	0.95	0.07	Positive
204	63	FR124_1	64	FR124_4	109	1.00	.	Boundary
205	65	FR305_11	67	FR305_21	34	0.00	0.00	.
206	65	FR305_11	71	FR353_1	20	0.00	0.00	.
207	65	FR305_11	72	FR353_2	20	0.00	0.00	.
208	65	FR305_11	73	FR353_3	20	0.00	0.00	.
209	66	FR305_12	68	FR305_22	35	0.00	0.00	.
210	67	FR305_21	71	FR353_1	20	0.00	0.00	.
211	67	FR305_21	72	FR353_2	20	0.00	0.00	.
212	67	FR305_21	73	FR353_3	20	0.00	0.00	.
213	69	FR354_1	70	FR354_2	76	0.83	0.11	Positive
214	71	FR353_1	72	FR353_2	24	0.00	0.00	.
215	71	FR353_1	73	FR353_3	23	0.00	0.00	.
216	72	FR353_2	73	FR353_3	23	0.00	0.00	.
217	74	FR399_1	75	FR399_2	85	0.70	0.10	Positive
218	74	FR399_1	76	FR399_3	88	0.52	0.14	Positive
219	75	FR399_2	76	FR399_3	86	1.00	.	Boundary

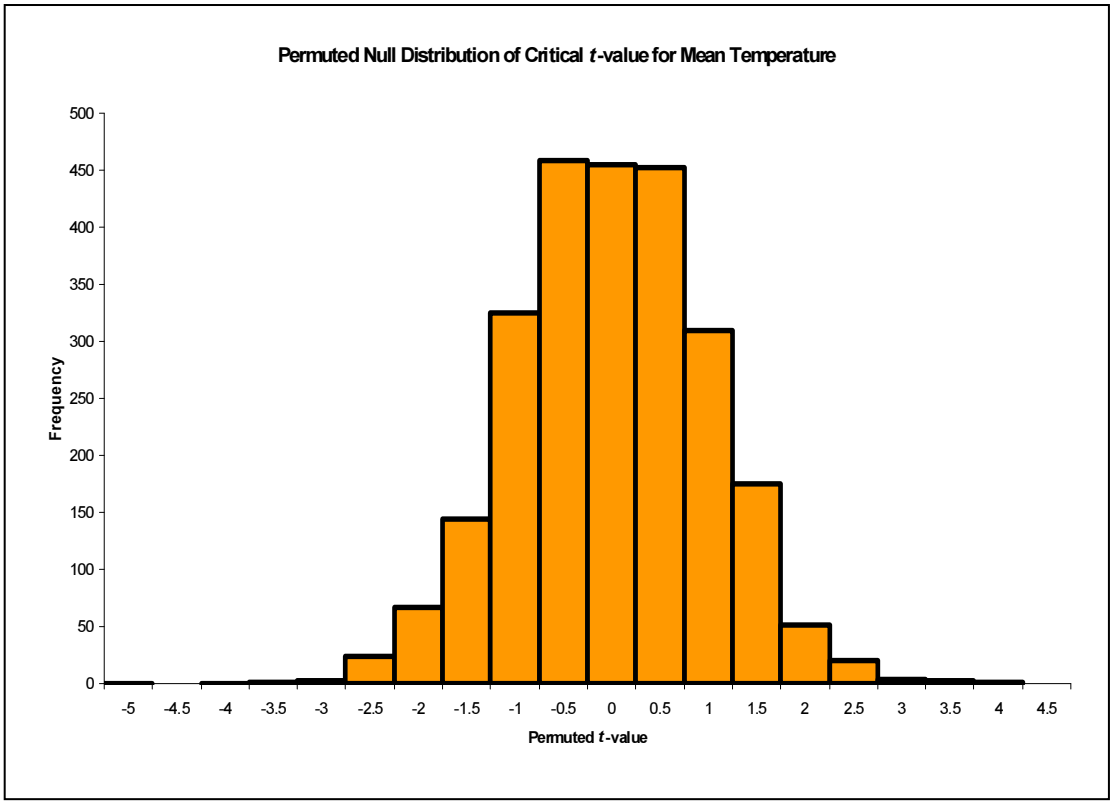
10. APPENDIX C – PERMUTED NULL DISRIBUTIONS FOR ENVIRONMENTAL VARIABLES USED IN MRM ANALYSIS

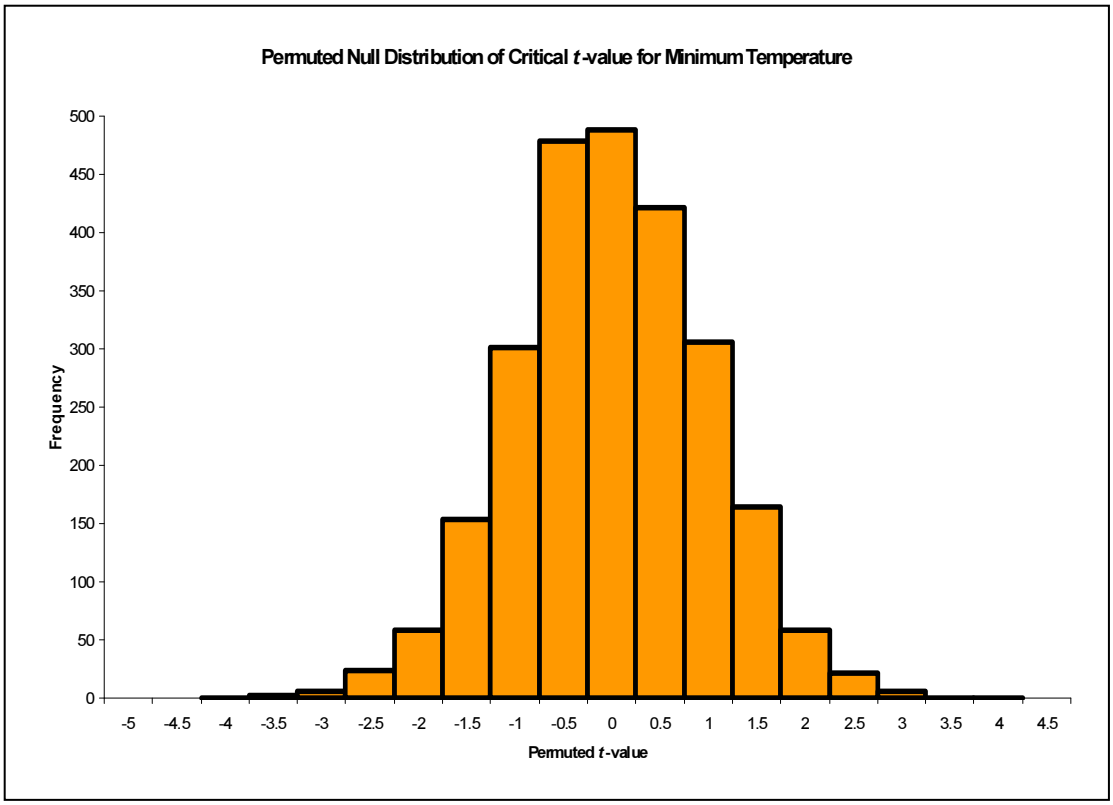


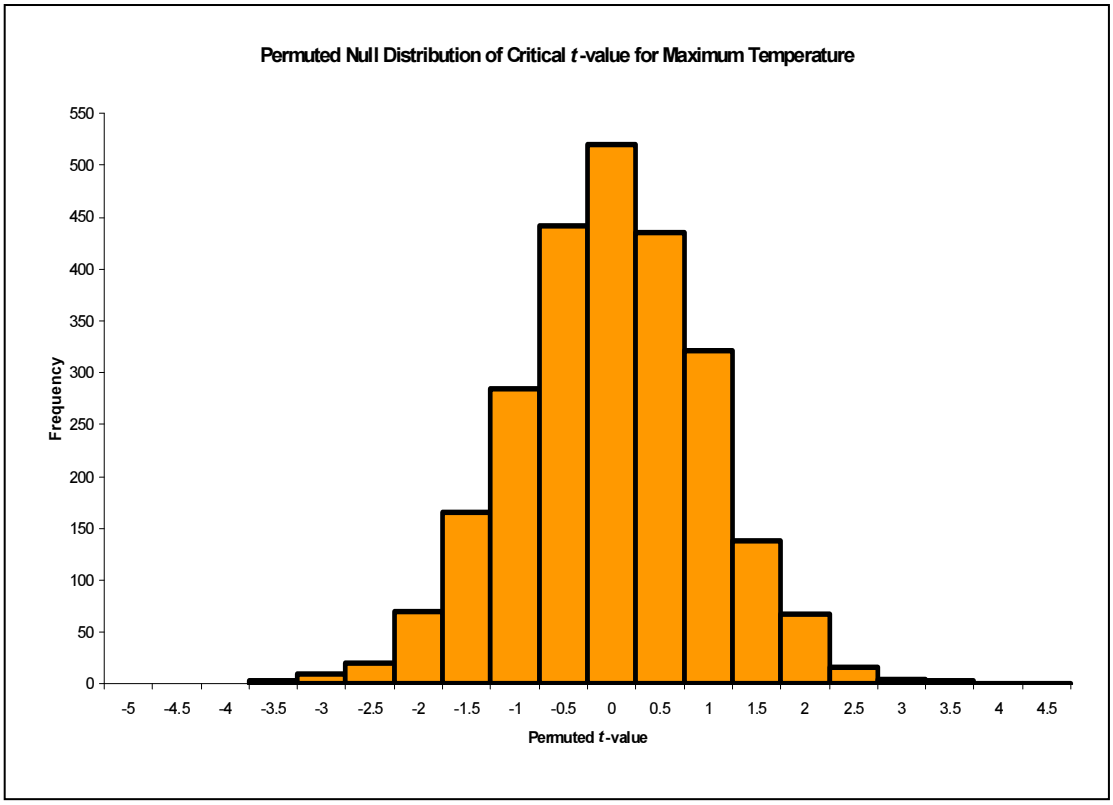


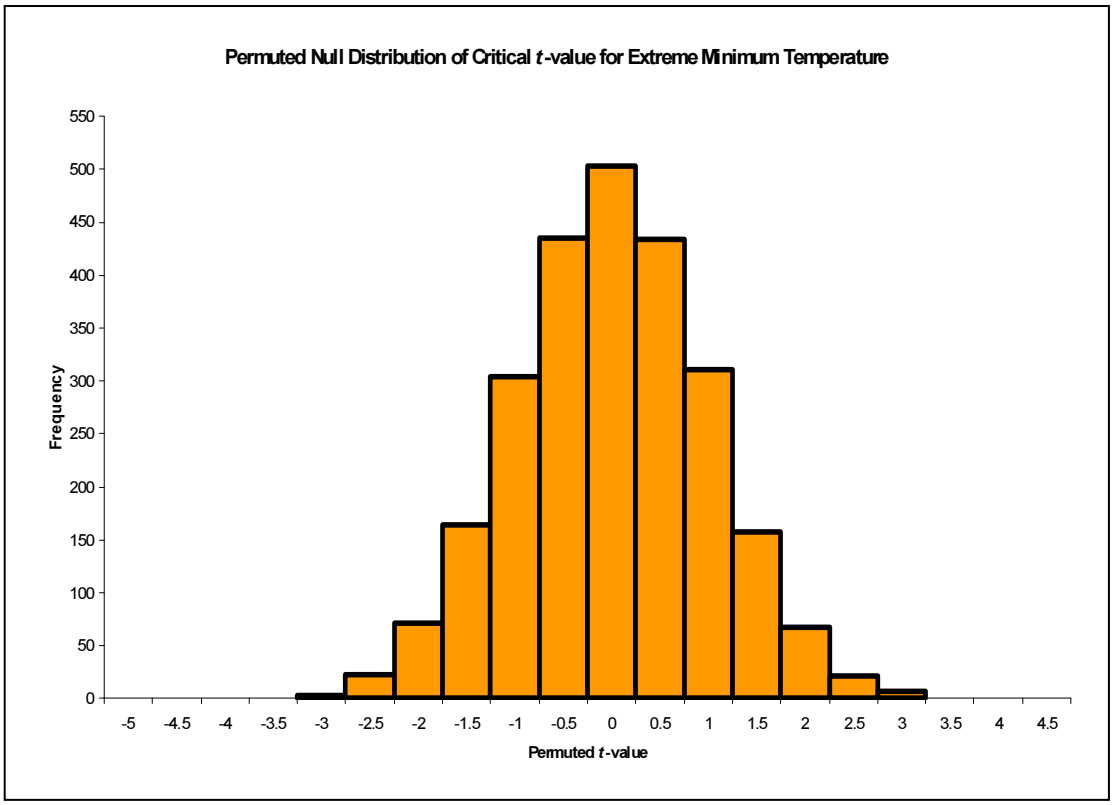

Permuted Null Distribution of Critical t-value for 24-hour Rainfall
Frequency
500
450
400
350
300
250
200
150
100
50
0
-5 -4.5 -4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4
Permuted t-value

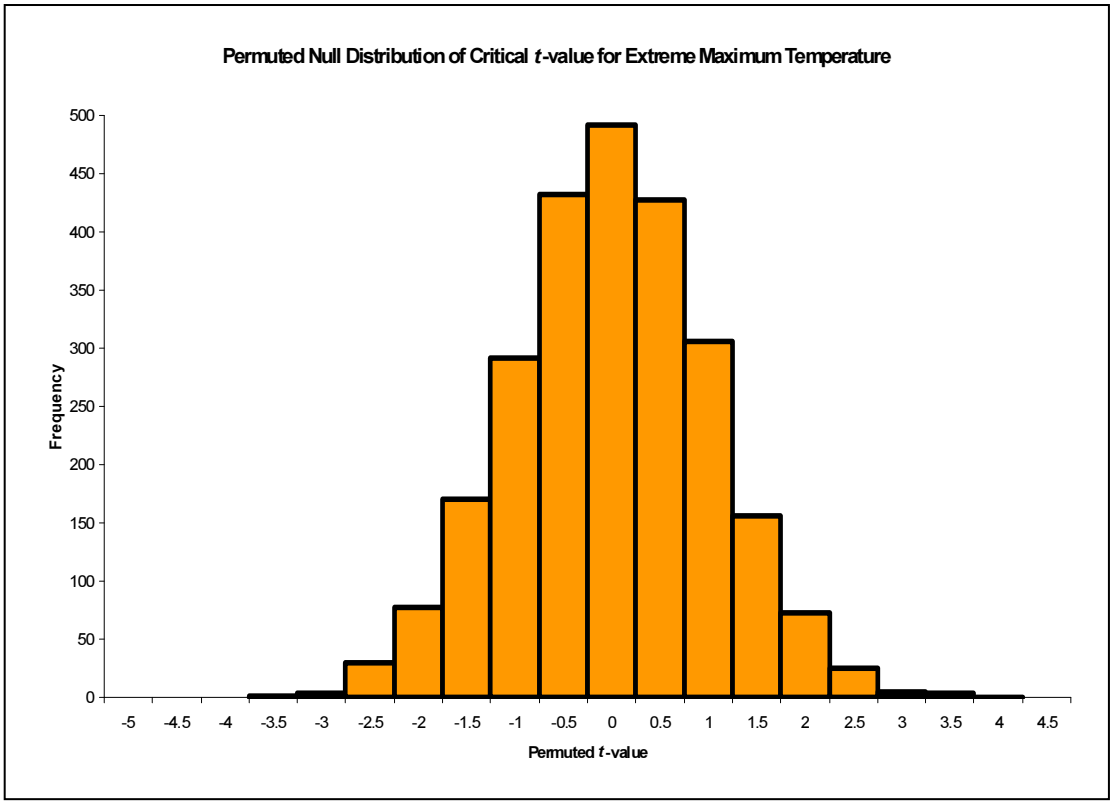

Permuted Null Distribution of Critical t-value for Vapour Pressure
Frequency
500
450
400
350
300
250
200
150
100
50
0
-5 -4.5 -4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Permuted t-value


Permuted Null Distribution of Critical t-value for Mean Temperature
Frequency
500
450
400
350
300
250
200
150
100
50
0
-5 -4.5 -4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Permuted t-value


Permuted Null Distribution of Critical t-value for Minimum Temperature
Frequency
500
450
400
350
300
250
200
150
100
50
0
-5 -4.5 -4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Permuted t-value


Permuted Null Distribution of Critical t-value for Maximum Temperature
Frequency
550
500
450
400
350
300
250
200
150
100
50
0
-5 -4.5 -4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Permuted t-value


Permuted Null Distribution of Critical t-value for Extreme Minimum Temperature
Frequency
550
500
450
400
350
300
250
200
150
100
50
0
-5 -4.5 -4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Permuted t-value


Permuted Null Distribution of Critical t-value for Extreme Maximum Temperature
Frequency
500
450
400
350
300
250
200
150
100
50
0
-5
-4.5
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
Permuted t-value

11. APPENDIX D – SAS OUTPUT FOR MRM ANALYSIS

1st GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	0.23966567	0.02662952	1.38	0.1911
Error	2201	42.44716680	0.01928540		
Corrected Total	2210	42.68683247			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.005615	476.6519	0.138872	0.029135

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Temp	1	0.01220390	0.01220390	0.63	0.4264
Rain	1	0.01310639	0.01310639	0.68	0.4098
MaxTemp	1	0.01032484	0.01032484	0.54	0.4644
MinTemp	1	0.00177034	0.00177034	0.09	0.7619
ExMaxTmp	1	0.03037181	0.03037181	1.57	0.2096
ExMinTmp	1	0.03602335	0.03602335	1.87	0.1719
VapPres	1	0.00821769	0.00821769	0.43	0.5140
Rain1day	1	0.04255674	0.04255674	2.21	0.1376
Altitude	1	0.05099359	0.05099359	2.64	0.1041

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0256712579	0.00777391	3.30	0.0010
Temp	0.0113579173	0.01427788	0.80	0.4264
Rain	0.0001809314	0.00021948	0.82	0.4098
MaxTemp	-.0062879130	0.00859368	-0.73	0.4644
MinTemp	-.0025491032	0.00841345	-0.30	0.7619
ExMaxTmp	0.0037440127	0.00298343	1.25	0.2096
ExMinTmp	-.0051483904	0.00376698	-1.37	0.1719
VapPres	-.0027055667	0.00414474	-0.65	0.5140
Rain1day	-.0009104066	0.00061287	-1.49	0.1376
Altitude	0.0000319961	0.00001968	1.63	0.1041

2nd GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	0.23789533	0.02973692	1.54	0.1374
Error	2202	42.44893714	0.01927745		
Corrected Total	2210	42.68683247			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.005573	476.5536	0.138843	0.029135

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Temp	1	0.02376609	0.02376609	1.23	0.2670
Rain	1	0.01469494	0.01469494	0.76	0.3827
MaxTemp	1	0.01068257	0.01068257	0.55	0.4567
ExMaxTmp	1	0.02908202	0.02908202	1.51	0.2195
ExMinTmp	1	0.05320534	0.05320534	2.76	0.0968
VapPres	1	0.01132276	0.01132276	0.59	0.4435
Rain1day	1	0.04408228	0.04408228	2.29	0.1306
Altitude	1	0.05178171	0.05178171	2.69	0.1014

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0253469235	0.00769827	3.29	0.0010
Temp	0.0075550685	0.00680431	1.11	0.2670
Rain	0.0001898513	0.00021745	0.87	0.3827
MaxTemp	-.0043971059	0.00590682	-0.74	0.4567
ExMaxTmp	0.0036387040	0.00296251	1.23	0.2195
ExMinTmp	-.0056418442	0.00339600	-1.66	0.0968
VapPres	-.0030523508	0.00398275	-0.77	0.4435
Rain1day	-.0009240691	0.00061108	-1.51	0.1306
Altitude	0.0000322198	0.00001966	1.64	0.1014

3rd GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	0.22721276	0.03245897	1.68	0.1084
Error	2203	42.45961970	0.01927355		
Corrected Total	2210	42.68683247			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.005323	476.5054	0.138829	0.029135

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Temp	1	0.01477198	0.01477198	0.77	0.3814
Rain	1	0.01642182	0.01642182	0.85	0.3561
ExMaxTmp	1	0.01955069	0.01955069	1.01	0.3140
ExMinTmp	1	0.04266161	0.04266161	2.21	0.1370
VapPres	1	0.01202442	0.01202442	0.62	0.4297
Rain1day	1	0.04760622	0.04760622	2.47	0.1162
Altitude	1	0.05771657	0.05771657	2.99	0.0837

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0247125149	0.00765018	3.23	0.0013
Temp	0.0033906216	0.00387294	0.88	0.3814
Rain	0.0002002797	0.00021697	0.92	0.3561
ExMaxTmp	0.0026990703	0.00267987	1.01	0.3140
ExMinTmp	-.0043966896	0.00295521	-1.49	0.1370
VapPres	-.0031440013	0.00398044	-0.79	0.4297
Rain1day	-.0009576707	0.00060935	-1.57	0.1162
Altitude	0.0000338137	0.00001954	1.73	0.0837

4th GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.21518834	0.03586472	1.86	0.0839
Error	2204	42.47164412	0.01927026		
Corrected Total	2210	42.68683247			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.005041	476.4647	0.138817	0.029135

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Temp	1	0.00345668	0.00345668	0.18	0.6719
Rain	1	0.01509248	0.01509248	0.78	0.3763
ExMaxTmp	1	0.03180495	0.03180495	1.65	0.1990
ExMinTmp	1	0.04425061	0.04425061	2.30	0.1298
Rain1day	1	0.04835292	0.04835292	2.51	0.1133
Altitude	1	0.04986307	0.04986307	2.59	0.1078

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0237698989	0.00755588	3.15	0.0017
Temp	0.0010664422	0.00251798	0.42	0.6719
Rain	0.0001917653	0.00021669	0.88	0.3763
ExMaxTmp	0.0033006134	0.00256916	1.28	0.1990
ExMinTmp	-.0044752833	0.00295328	-1.52	0.1298
Rain1day	-.0009650389	0.00060922	-1.58	0.1133
Altitude	0.0000308404	0.00001917	1.61	0.1078

5th GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	0.21173166	0.04234633	2.20	0.0519
Error	2205	42.47510080	0.01926308		
Corrected Total	2210	42.68683247			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.004960	476.3760	0.138792	0.029135

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Rain	1	0.01476319	0.01476319	0.77	0.3814
ExMaxTmp	1	0.05500203	0.05500203	2.86	0.0912
ExMinTmp	1	0.04156593	0.04156593	2.16	0.1420
Rain1day	1	0.04826567	0.04826567	2.51	0.1136
Altitude	1	0.05685230	0.05685230	2.95	0.0859

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0239934419	0.00753602	3.18	0.0015
Rain	0.0001896095	0.00021659	0.88	0.3814
ExMaxTmp	0.0038181301	0.00225956	1.69	0.0912
ExMinTmp	-.0039743115	0.00270555	-1.47	0.1420
Rain1day	-.0009641623	0.00060911	-1.58	0.1136
Altitude	0.0000323538	0.00001883	1.72	0.0859

6th GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.19696848	0.04924212	2.56	0.0371
Error	2206	42.48986399	0.01926104		
Corrected Total	2210	42.68683247			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.004614	476.3508	0.138784	0.029135

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ExMaxTmp	1	0.06671718	0.06671718	3.46	0.0629
ExMinTmp	1	0.04565990	0.04565990	2.37	0.1238
Rain1day	1	0.03534139	0.03534139	1.83	0.1757
Altitude	1	0.05590590	0.05590590	2.90	0.0886

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0252690009	0.00739343	3.42	0.0006
ExMaxTmp	0.0041467159	0.00222805	1.86	0.0629
ExMinTmp	-.0041535003	0.00269766	-1.54	0.1238
Rain1day	-.0005942787	0.00043872	-1.35	0.1757
Altitude	0.0000320790	0.00001883	1.70	0.0886

7th GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.16021989	0.05340663	2.85	0.0362
Error	2274	42.61991407	0.01874227		
Corrected Total	2277	42.78013396			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.003745	481.7403	0.136902	0.028418

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ExMaxTmp	1	0.05295960	0.05295960	2.83	0.0929
ExMinTmp	1	0.04085415	0.04085415	2.18	0.1400
Altitude	1	0.06197171	0.06197171	3.31	0.0691

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0197424412	0.00641398	3.08	0.0021
ExMaxTmp	0.0036024896	0.00214309	1.68	0.0929
ExMinTmp	-.0038656664	0.00261829	-1.48	0.1400
Altitude	0.0000334429	0.00001839	1.82	0.0691

8th GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	0.11936575	0.05968287	3.18	0.0417
Error	2275	42.66076821	0.01875199		
Corrected Total	2277	42.78013396			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.002790	481.8651	0.136938	0.028418

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ExMaxTmp	1	0.05049450	0.05049450	2.69	0.1009
Altitude	1	0.04225354	0.04225354	2.25	0.1335

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0157090060	0.00580459	2.71	0.0069
ExMaxTmp	0.0035163442	0.00214286	1.64	0.1009
Altitude	0.0000267670	0.00001783	1.50	0.1335

9th GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.07711220	0.07711220	4.11	0.0427
Error	2276	42.70302176	0.01876231		
Corrected Total	2277	42.78013396			

R-Square	Coeff Var	Root MSE	SiteDist Mean
0.001803	481.9978	0.136976	0.028418

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ExMaxTmp	1	0.07711220	0.07711220	4.11	0.0427

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0205621665	0.00482216	4.26	<.0001
ExMaxTmp	0.0042354354	0.00208920	2.03	0.042

12. APPENDIX E – RESULTS OF CLUSTERING USING THRESHOLD ANALYSIS TECHNIQUE

Alpha	Cluster number	Cluster members
0.04	1	58
0.04	2	59
0.05	1	31
0.05	2	47
0.05	3	58
0.05	4	59
0.06	1	31
0.06	2	47
0.06	3	58
0.06	4	59
0.06	5	61
0.06	6	62
0.07	1	10
0.07	2	31
0.07	3	32
0.07	4	47
0.07	5	58
0.07	6	59
0.07	7	61
0.07	8	62
0.09	1	10,58
0.09	2	31
0.09	3	32,59
0.09	4	34,47
0.09	5	53
0.09	6	60
0.09	7	61
0.09	8	62
0.11	1	10,58
0.11	2	31
0.11	3	32,59
0.11	4	34,47
0.11	5	51
0.11	6	52
0.11	7	53
0.11	8	60
0.11	9	61
0.11	10	62
0.12	1	10,58
0.12	2	31
0.12	3	32,59
0.12	4	34,47,76
0.12	5	51
0.12	6	52
0.12	7	53
0.12	8	60
0.12	9	61
0.12	10	62
0.13	1	10,34,47,58,76
0.13	2	31,32,59
0.13	3	41
0.13	4	44
0.13	5	51
0.13	6	52
0.13	7	53
0.13	8	60
0.13	9	61
0.13	10	62
0.14	1	10,34,47,58,76
0.14	2	31,32,59
0.14	3	35
0.14	4	36
0.14	5	41
0.14	6	44
0.14	7	51
0.14	8	52
0.14	9	53
0.14	10	60
0.14	11	61
0.14	12	62
0.16	1	10,34,47,58,76
0.16	2	31,32,59
0.16	3	35
0.16	4	36
0.16	5	41
0.16	6	44
0.16	7	51
0.16	8	52
0.16	9	53
0.16	10	60
0.16	11	61
0.16	12	62
0.16	13	63
0.16	14	7
0.17	1	10,34,47,58,76
0.17	2	2
0.17	3	29
0.17	4	31,32,59
0.17	5	35
0.17	6	36
0.17	7	41
0.17	8	44
0.17	9	51
0.17	10	52
0.17	11	53
0.17	12	60
0.17	13	61
0.17	14	62
0.17	15	63
0.17	16	69
0.17	17	7
0.17	18	70
0.19	1	10,34,47,58,76
0.19	2	16
0.19	3	18
0.19	4	2
0.19	5	27
0.19	6	28
0.19	7	29
0.19	8	31,32,59
0.19	9	35
0.19	10	36
0.19	11	41
0.19	12	44
0.19	13	51
0.19	14	52

Alpha	Cluster number	Cluster members
0.19	15	53
0.19	16	60
0.19	17	61
0.19	18	62
0.19	19	63
0.19	20	69
0.19	21	7
0.19	22	70
0.20	1	10,34,47,58,76
0.20	2	16
0.20	3	18
0.20	4	2
0.20	5	27
0.20	6	28
0.20	7	29
0.20	8	31,32,48,59
0.20	9	35
0.20	10	36
0.20	11	41
0.20	12	44
0.20	13	51
0.20	14	52
0.20	15	53
0.20	16	60
0.20	17	61
0.20	18	62
0.20	19	63
0.20	20	69
0.20	21	7
0.20	22	70
0.21	1	10,34,47,58,75,76
0.21	2	16
0.21	3	18
0.21	4	2
0.21	5	27
0.21	6	28
0.21	7	29
0.21	8	31,32,48,59
0.21	9	35
0.21	10	36
0.21	11	41
0.21	12	44
0.21	13	51
0.21	14	52
0.21	15	53
0.21	16	60
0.21	17	61
0.21	18	62
0.21	19	63,64
0.21	20	69
0.21	21	7
0.21	22	70
0.22	1	10,34,47,58,75,76
0.22	2	16
0.22	3	18
0.22	4	2
0.22	5	27
0.22	6	28

Alpha	Cluster number	Cluster members
0.22	7	29
0.22	8	31,32,48,59
0.22	9	35
0.22	10	36
0.22	11	41
0.22	12	44
0.22	13	50,51
0.22	14	52
0.22	15	53
0.22	16	60
0.22	17	61
0.22	18	62
0.22	19	63,64
0.22	20	69
0.22	21	7
0.22	22	70
0.24	1	10,34,47,58,75,76
0.24	2	15,18
0.24	3	16
0.24	4	2
0.24	5	24
0.24	6	25
0.24	7	27
0.24	8	28
0.24	9	29
0.24	10	31,32,48,59
0.24	11	35
0.24	12	36
0.24	13	41
0.24	14	44
0.24	15	50,51
0.24	16	52
0.24	17	53
0.24	18	60
0.24	19	61
0.24	20	62
0.24	21	63,64
0.24	22	69
0.24	23	7
0.24	24	70
0.25	1	10,34,47,58,75,76
0.25	2	15,18
0.25	3	16
0.25	4	2
0.25	5	24
0.25	6	25
0.25	7	27
0.25	8	28
0.25	9	29
0.25	10	30,31,32,48,59
0.25	11	35
0.25	12	36
0.25	13	41
0.25	14	44
0.25	15	50,51
0.25	16	52
0.25	17	53
0.25	18	60

Alpha	Cluster number	Cluster members
0.25	19	61
0.25	20	62
0.25	21	63,64
0.25	22	69
0.25	23	7
0.25	24	70
0.26	1	1
0.26	2	10,30,31,32,34,47,48,58,59,75,76
0.26	3	15,18
0.26	4	16,17
0.26	5	2
0.26	6	24
0.26	7	25
0.26	8	27
0.26	9	28
0.26	10	29
0.26	11	3
0.26	12	35
0.26	13	36
0.26	14	41
0.26	15	44
0.26	16	50,51
0.26	17	52
0.26	18	53
0.26	19	6,63,64
0.26	20	60
0.26	21	61
0.26	22	62
0.26	23	69
0.26	24	7
0.26	25	70
0.27	1	1
0.27	2	10,30,31,32,33,34,47,48,58,59,75,76
0.27	3	15,18
0.27	4	16,17
0.27	5	2
0.27	6	24
0.27	7	25
0.27	8	27
0.27	9	28
0.27	10	29
0.27	11	3
0.27	12	35
0.27	13	36
0.27	14	41
0.27	15	44
0.27	16	50,51
0.27	17	52
0.27	18	53
0.27	19	6,63,64
0.27	20	60
0.27	21	61
0.27	22	62
0.27	23	69
0.27	24	7
0.27	25	70
0.28	1	1
0.28	2	10,30,31,32,33,34,47,48,58,59,75,76

Alpha	Cluster number	Cluster members
0.28	3	15,18
0.28	4	16,17
0.28	5	2
0.28	6	20
0.28	7	21
0.28	8	24
0.28	9	25
0.28	10	27
0.28	11	28
0.28	12	29
0.28	13	3
0.28	14	35
0.28	15	36
0.28	16	41
0.28	17	44
0.28	18	50,51
0.28	19	52
0.28	20	53
0.28	21	6,63,64
0.28	22	60
0.28	23	61
0.28	24	62
0.28	25	69
0.28	26	7
0.28	27	70
0.30	1	1
0.30	2	10,30,31,32,33,34,47,48,58,59,75,76
0.30	3	12
0.30	4	13
0.30	5	15,18
0.30	6	16,17
0.30	7	2
0.30	8	20
0.30	9	21
0.30	10	24
0.30	11	25
0.30	12	27
0.30	13	28
0.30	14	29
0.30	15	3
0.30	16	35
0.30	17	36
0.30	18	41
0.30	19	44
0.30	20	50,51
0.30	21	52
0.30	22	53
0.30	23	6,63,64
0.30	24	60
0.30	25	61
0.30	26	62
0.30	27	69
0.30	28	7
0.30	29	70
0.31	1	1,3,10,30,31,32,33,34,47,48,58,59,74,75,76
0.31	2	12
0.31	3	13
0.31	4	15,18

Alpha	Cluster number	Cluster members
0.31	5	16,17
0.31	6	2
0.31	7	20
0.31	8	21
0.31	9	24
0.31	10	25
0.31	11	27
0.31	12	28
0.31	13	29
0.31	14	35
0.31	15	36
0.31	16	41
0.31	17	44
0.31	18	50,51
0.31	19	52
0.31	20	53
0.31	21	6,63,64
0.31	22	60
0.31	23	61
0.31	24	62
0.31	25	69
0.31	26	7
0.31	27	70
0.32	1	1,3,10,30,31,32,33,34,47,48,58,59,74,75,76
0.32	2	12
0.32	3	13
0.32	4	15,16,17,18
0.32	5	2
0.32	6	20
0.32	7	21
0.32	8	24
0.32	9	25
0.32	10	26,28
0.32	11	27
0.32	12	29
0.32	13	35
0.32	14	36
0.32	15	41
0.32	16	44
0.32	17	50,51
0.32	18	52
0.32	19	53
0.32	20	6,63,64
0.32	21	60
0.32	22	61
0.32	23	62
0.32	24	69
0.32	25	7
0.32	26	70
0.33	1	1,2,3,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.33	2	12
0.33	3	13
0.33	4	15,16,17,18
0.33	5	20
0.33	6	21
0.33	7	24
0.33	8	25
0.33	9	26,28

Alpha	Cluster number	Cluster members
0.33	10	27
0.33	11	35
0.33	12	36
0.33	13	41,46
0.33	14	44
0.33	15	50,51
0.33	16	52
0.33	17	53
0.33	18	6,63,64
0.33	19	60
0.33	20	61
0.33	21	62
0.33	22	69
0.33	23	7
0.33	24	70
0.36	1	1,2,3,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.36	2	12
0.36	3	13
0.36	4	15,16,17,18
0.36	5	20
0.36	6	21
0.36	7	24
0.36	8	25
0.36	9	26,28
0.36	10	27
0.36	11	35
0.36	12	36
0.36	13	41,46
0.36	14	44
0.36	15	50,51,52
0.36	16	53
0.36	17	6,63,64
0.36	18	60
0.36	19	61
0.36	20	62
0.36	21	69
0.36	22	7
0.36	23	70
0.37	1	1,2,3,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.37	2	12
0.37	3	13
0.37	4	15,16,17,18,19
0.37	5	20
0.37	6	21
0.37	7	24
0.37	8	25
0.37	9	26,27,28
0.37	10	35
0.37	11	36
0.37	12	41,46
0.37	13	44
0.37	14	50,51,52
0.37	15	53
0.37	16	6,63,64
0.37	17	60
0.37	18	61
0.37	19	62
0.37	20	69

Alpha	Cluster number	Cluster members
0.37	21	7
0.37	22	70
0.39	1	1,2,3,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.39	2	12
0.39	3	13
0.39	4	15,16,17,18,19
0.39	5	20
0.39	6	21
0.39	7	23,24
0.39	8	25
0.39	9	26,27,28
0.39	10	35
0.39	11	36
0.39	12	41,46
0.39	13	44
0.39	14	50,51,52
0.39	15	53
0.39	16	6,63,64
0.39	17	60
0.39	18	61
0.39	19	62
0.39	20	69
0.39	21	7
0.39	22	70
0.40	1	1,2,3,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.40	2	12
0.40	3	13
0.40	4	15,16,17,18,19
0.40	5	20
0.40	6	21
0.40	7	23,24
0.40	8	25
0.40	9	26,27,28
0.40	10	35
0.40	11	36
0.40	12	41,42,46
0.40	13	44
0.40	14	50,51,52
0.40	15	53
0.40	16	6,63,64
0.40	17	60
0.40	18	61
0.40	19	62
0.40	20	69
0.40	21	7
0.40	22	70
0.41	1	1,2,3,5,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.41	2	12
0.41	3	13,14
0.41	4	15,16,17,18,19
0.41	5	20
0.41	6	21
0.41	7	23,24
0.41	8	25
0.41	9	26,27,28
0.41	10	35
0.41	11	36
0.41	12	41,42,46

Alpha	Cluster number	Cluster members
0.41	13	44
0.41	14	50,51,52
0.41	15	53
0.41	16	6,63,64
0.41	17	60
0.41	18	61
0.41	19	62
0.41	20	69
0.41	21	7
0.41	22	70
0.42	1	1,2,3,5,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.42	2	11,12
0.42	3	13,14
0.42	4	15,16,17,18,19,50,51,52
0.42	5	20
0.42	6	21
0.42	7	23,24
0.42	8	25
0.42	9	26,27,28
0.42	10	35
0.42	11	36
0.42	12	41,42,46
0.42	13	44
0.42	14	53
0.42	15	6,63,64
0.42	16	60
0.42	17	61
0.42	18	62
0.42	19	69
0.42	20	7
0.42	21	70
0.43	1	1,2,3,5,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.43	2	11,12
0.43	3	13,14
0.43	4	15,16,17,18,19,50,51,52
0.43	5	20
0.43	6	21
0.43	7	23,24,25
0.43	8	26,27,28
0.43	9	35
0.43	10	36
0.43	11	41,42,46
0.43	12	44
0.43	13	53
0.43	14	6,63,64
0.43	15	60
0.43	16	61
0.43	17	62
0.43	18	69
0.43	19	7
0.43	20	70
0.44	1	1,2,3,4,5,10,29,30,31,32,33,34,47,48,58,59,74,75,76
0.44	2	11,12
0.44	3	13,14
0.44	4	15,16,17,18,19,50,51,52
0.44	5	20
0.44	6	21
0.44	7	23,24,25

Alpha	Cluster number	Cluster members
0.44	8	26,27,28
0.44	9	35
0.44	10	36
0.44	11	41,42,46
0.44	12	44
0.44	13	53
0.44	14	54,60
0.44	15	6,63,64
0.44	16	61
0.44	17	62
0.44	18	69
0.44	19	7
0.44	20	70
0.45	1	1,2,3,4,5,10,29,30,31,32,33,34,47,48,55,58,59,74,75,76
0.45	2	11,12
0.45	3	13,14
0.45	4	15,16,17,18,19,50,51,52
0.45	5	20
0.45	6	21
0.45	7	23,24,25
0.45	8	26,27,28
0.45	9	35
0.45	10	36
0.45	11	41,42,46
0.45	12	44
0.45	13	53
0.45	14	54,60
0.45	15	6,63,64
0.45	16	61
0.45	17	62
0.45	18	69
0.45	19	7
0.45	20	70
0.47	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,58,59,74,75,76
0.47	2	15,16,17,18,19,50,51,52
0.47	3	20
0.47	4	21
0.47	5	23,24,25
0.47	6	26,27,28
0.47	7	35
0.47	8	36
0.47	9	41,42,46
0.47	10	44
0.47	11	53
0.47	12	54,60
0.47	13	6,7,63,64
0.47	14	61
0.47	15	62
0.47	16	69
0.47	17	70
0.48	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.48	2	15,16,17,18,19,50,51,52
0.48	3	20
0.48	4	21
0.48	5	23,24,25
0.48	6	26,27,28
0.48	7	35
0.48	8	36

Alpha	Cluster number	Cluster members
0.48	9	41,42,46
0.48	10	44
0.48	11	53
0.48	12	54,60
0.48	13	6,7,63,64
0.48	14	61
0.48	15	62
0.48	16	69
0.48	17	70
0.50	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.50	2	15,16,17,18,19,50,51,52
0.50	3	20
0.50	4	21,22
0.50	5	23,24,25
0.50	6	26,27,28
0.50	7	35
0.50	8	36
0.50	9	41,42,46
0.50	10	44
0.50	11	53
0.50	12	54,60
0.50	13	6,7,63,64
0.50	14	61
0.50	15	62
0.50	16	69
0.50	17	70
0.52	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.52	2	15,16,17,18,19,50,51,52
0.52	3	20,21,22
0.52	4	23,24,25
0.52	5	26,27,28
0.52	6	35
0.52	7	36
0.52	8	41,42,46
0.52	9	44
0.52	10	53
0.52	11	54,60
0.52	12	6,7,63,64
0.52	13	61
0.52	14	62
0.52	15	69
0.52	16	70
0.56	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.56	2	15,16,17,18,19,50,51,52
0.56	3	20,21,22
0.56	4	23,24,25
0.56	5	26,27,28
0.56	6	35
0.56	7	36
0.56	8	37,44
0.56	9	41,42,46
0.56	10	53
0.56	11	54,60
0.56	12	6,7,63,64
0.56	13	61
0.56	14	62
0.56	15	69
0.56	16	70

Alpha	Cluster number	Cluster members
0.57	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.57	2	15,16,17,18,19,50,51,52
0.57	3	20,21,22
0.57	4	23,24,25
0.57	5	26,27,28
0.57	6	35
0.57	7	36
0.57	8	37,44
0.57	9	41,42,46
0.57	10	53,54,60
0.57	11	6,7,63,64
0.57	12	61
0.57	13	62
0.57	14	69
0.57	15	70
0.63	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.63	2	15,16,17,18,19,50,51,52
0.63	3	23,24,25
0.63	4	26,27,28
0.63	5	35
0.63	6	36
0.63	7	37,44
0.63	8	41,42,46
0.63	9	53,54,60
0.63	10	6,7,20,21,22,63,64
0.63	11	61
0.63	12	62
0.63	13	69
0.63	14	70
0.65	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.65	2	15,16,17,18,19,50,51,52
0.65	3	23,24,25
0.65	4	26,27,28
0.65	5	35
0.65	6	36
0.65	7	37,44
0.65	8	38,41,42,46
0.65	9	53,54,60
0.65	10	6,7,20,21,22,63,64
0.65	11	61
0.65	12	62
0.65	13	69
0.65	14	70
0.72	1	1,2,3,4,5,10,11,12,13,14,29,30,31,32,33,34,47,48,55,56,58,59,74,75,76
0.72	2	15,16,17,18,19,50,51,52
0.72	3	23,24,25
0.72	4	26,27,28
0.72	5	35
0.72	6	36
0.72	7	37,38,41,42,44,46
0.72	8	53,54,60
0.72	9	6,7,20,21,22,63,64
0.72	10	61
0.72	11	62
0.72	12	69
0.72	13	70